



# Méthode de détection et de suivi multi-piétons multi-capteurs embarquée sur un véhicule routier: application à un environnement urbain

Samuel Gidel

## ► To cite this version:

Samuel Gidel. Méthode de détection et de suivi multi-piétons multi-capteurs embarquée sur un véhicule routier: application à un environnement urbain. Automatique / Robotique. Université Blaise Pascal - Clermont-Ferrand II, 2010. Français. NNT: 2010CLF22028 . tel-00719262

**HAL Id: tel-00719262**

**<https://theses.hal.science/tel-00719262>**

Submitted on 19 Jul 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Blaise Pascal - Clermont II

Ecole Doctorale

Sciences Pour L'Ingénieur De Clermont-Ferrand

Thèse

présentée par

Samuel GIDEL

Ingénieur du C.U.S.T

pour obtenir le grade de

Docteur d'Université

(Spécialité : Vision pour la robotique)

---

Méthodes de détection et de suivi multi-piétons  
multi-capteurs embarquées sur un véhicule  
routier : application à un environnement  
urbain

---

Soutenue publiquement le **29 avril 2010** devant le jury :

Monsieur D. AUBERT	Président
Madame V. BERGE CHERFAOUI	Rapporteur
Monsieur M. DEVY	Rapporteur
Monsieur R. REYNAUD	Examineur
Monsieur C. BLANC	Examineur
Monsieur T. CHATEAU	Examineur
Monsieur P. CHECCHIN	Examineur
Monsieur L. TRASSOUDAIN	Directeur de thèse

---

*à Adelphine,  
Noémie,  
mes parents.*



---

---

# Remerciement

Les travaux présentés dans cette thèse ont été effectués au sein de l'antenne de Montluçon du Laboratoire LASMEA, Unité Mixte de recherche Université Blaise Pascal - CNRS de Clermont-Ferrand.

Je tiens à remercier Didier Aubert pour avoir accepté de présider ce jury de thèse. Je remercie Véronique Berge Cherfaoui et Michel Devy pour avoir apporté leurs points de vue sur ces travaux en acceptant d'en être les rapporteurs. Je suis également reconnaissant à Roger Raynaud pour avoir accepté de juger ce travail.

Je remercie Laurent Trassoudaine, mon directeur de thèse pour m'avoir accueilli dans son équipe de recherche. Un grand merci à Paul Checchin et Christophe Blanc mes encadrants montluçonnais qui par leurs conseils, leurs disponibilités et leurs compétences m'ont permis de mener à bien le travail présenté dans ce manuscrit.

Je suis particulièrement reconnaissant à Thierry Chateau pour m'avoir proposé ce sujet de thèse puis pour son soutien, son enthousiasme et le grand intérêt qu'il a porté à mes travaux malgré la distance séparant Montluçon de Clermont-Ferrand.

Merci à tous ceux qui m'ont entouré et aidé durant ces trois années. Merci en particulier à Laurent Malaterre et François Marmoiton pour leurs disponibilités lors des expérimentations effectuées sur les véhicules du LASMEA. Je voudrais remercier très spécialement Michelle et Adrien Belanger pour les heures qu'ils ont consacrées à la relecture de cette thèse.

Enfin, merci à Adeline et à mes parents pour leur soutien sans faille tout au long de ce travail.



---

# Résumé

Les travaux présentés dans cette thèse ont pour cadre la vision par ordinateur et concernent la détection et le suivi de piéton se trouvant sur la trajectoire d'un véhicule routier circulant en milieu urbain. Dans ce type d'environnement complexe, une des difficultés majeures est la capacité à discerner les piétons des nombreux autres obstacles situés sur la chaussée. Un autre point essentiel est de pouvoir les suivre afin de prédire leur déplacement et ainsi le cas échéant éviter le contact avec le véhicule. D'autres contraintes s'ajoutent dans le contexte industriel des véhicules routiers intelligents. Il est nécessaire de proposer des algorithmes robustes temps réel avec des capteurs les moins chers possible.

Dans une première partie, mon travail a consisté à mettre au point une méthode de détection, d'identification puis de suivi de piéton en environnement extérieur à partir d'un unique capteur laser quatre plans. L'extraction d'« objets piétons » et la fusion des quatre plans laser reposent tous deux, sur une méthode non paramétrique à noyau aussi appelée « fenêtres de parzen ». Cet estimateur permet dans un premier temps, d'approximer la fonction de vraisemblance d'un relevé d'impacts laser en fonction des caractéristiques géométriques d'un piéton. Dans un second temps, cet estimateur est utilisé pour calculer la probabilité qu'un piéton soit situé dans les quatre plans laser. Afin de caractériser au mieux la trajectoire complexe d'un piéton, le pistage repose sur l'utilisation classique d'un filtre à particules.

Malheureusement un système de détection de piéton fondé uniquement sur un capteur laser reste insuffisant du point de vue de la performance. En effet, les limitations intrinsèques de ce capteur (absence d'information sur la hauteur, le contour, ou la couleur des objets), sa sensibilité aux conditions atmosphériques comme la pluie ou le brouillard, a conduit à réaliser une solution multisensorielle permettant de combiner judicieusement les informations fournies par un capteur laser et vidéo. Cette fusion repose sur le développement d'une méthode non paramétrique d'association de données permettant de conserver toute l'information contenue dans les mesures envoyées par les capteurs laser et vidéo.

Les performances des différents algorithmes proposés ont été caractérisées et commentées à partir de données réelles obtenues sur de multiples scénarios enregistrés à bord du véhicule expérimental du LASMEA et de Renault, avec qui nous avons collaboré dans le cadre du projet ANR LOVE.

**Mots clés :** Détection et suivi de piéton, LIDAR, véhicule intelligent, filtre à particules, méthodes non paramétriques, fusion de données laser et vidéo.



---

# Abstract

The research work presented for this thesis is based on computer vision and deals with the detection and tracking of pedestrians located on a vehicle's trajectory in an urban environment. In such a complex environment, one of the major difficulties is the ability to detect pedestrians among numerous other obstacles on the road ahead of the vehicle. Another key point is the ability to follow pedestrians so as to predict their trajectory and thus, if necessary, to avoid any collision with the vehicle. Several other constraints have to be faced in the industrial context of smart vehicles. Therefore some robust, real-time algorithms are required along with the cheapest possible sensors.

At first my work consisted in developing an appropriate method to detect, then identify and track pedestrians in an outdoor environment from a single four-plane laser sensor. Pedestrian extraction and the merging of the four laser planes are both based on a non-parametric kernel method, also called "Parzen Windows". Initially, this estimator is used to approximate the likelihood function of the impact record in the laser image according to a pedestrian's geometrical characteristics. Secondly this estimator is used to calculate the likelihood that a pedestrian should be located within the four laser planes. Finally, to best characterize the complex trajectory of a pedestrian, the tracking process is based on the traditional particle filter.

Unfortunately a pedestrian detection system which relies only on a laser sensor remains unsatisfactory as far as performance is concerned. Indeed, the inherent limitations of this sensor (no information about height, the outline or the color of the objects), as well as its sensitivity to such atmospheric conditions as rain or fog, make it necessary to resort to a multisensorial solution which allows to effectively combine the information provided by the laser and video sensors. This fusion methods is based on the development of a non-parametric method for data association, which allows to keep all the information contained in the measurements sent by the laser and video sensors.

The performance of each proposed algorithm was characterized and reviewed, using real data obtained from numerous recording on board the LASMEA and Renault test vehicle ; Renault being the French vehicle manufacturer with whom we collaborate on our ANR LOVE project.

## Keys words

Pedestrian detection, Pedestrian tracking, LIDAR, intelligent vehicle, particle filter, non-parametric methods, laser and video data fusion.



---

# Table des matières

Remerciement . . . . .	iii
Résumé . . . . .	v
Abstract . . . . .	vii
Table des matières . . . . .	ix
Table des figures . . . . .	xiii
Liste des tableaux . . . . .	xxiii
Liste des algorithmes . . . . .	xxv
Notations et Conventions . . . . .	xxvii
<b>1 Introduction générale</b>	<b>1</b>
1.1 Généralités sur la sécurité routière des vulnérables . . . . .	1
1.1.1 Les usagers de la route : la perspective européenne . . . . .	1
1.1.2 Les principaux usagers concernés . . . . .	2
1.2 La protection des piétons dans les véhicules routiers . . . . .	3
1.2.1 Introduction . . . . .	3
1.2.2 EuroNCAP . . . . .	4
1.3 Les projets traitant la protection des piétons depuis 2002 . . . . .	5
1.3.1 SAVE-U : Capteurs et architecture système pour la protection des vulnérables . . . . .	6
1.3.2 PReVENT : Applications de sécurité active et préventive . . . . .	7
1.3.3 Le projet LOVE . . . . .	7
1.4 Contributions de la thèse . . . . .	9
1.4.1 Formulation d'une méthode de détection d'« objets-piétons » exploitant des données laser . . . . .	10
1.4.2 Formulation d'une méthode d'association de données pour filtre particulière . . . . .	11
1.5 Organisation du document . . . . .	12
<b>2 Perception par télémétrie laser</b>	<b>15</b>
2.1 Principe de la télémétrie laser à balayage . . . . .	15
2.1.1 Généralité . . . . .	15
2.1.2 Du capteur à l'image . . . . .	16
2.1.3 Bilan des capteurs LIDAR exploitables . . . . .	17
2.1.4 Capteur retenu . . . . .	20
2.1.5 Conclusion . . . . .	22
2.2 Méthodes de détection/reconnaissance d'obstacles . . . . .	23
2.2.1 Méthodes directes . . . . .	23
2.2.2 Méthodes inverses . . . . .	25



2.3	L'extraction de primitives piétons . . . . .	26
2.3.1	Les problèmes de la détection des piétons . . . . .	26
2.3.2	Etat de l'art des méthodes existantes . . . . .	30
2.3.3	De l'idéal à la réalité . . . . .	40
2.4	Conclusion . . . . .	43
<b>3</b>	<b>Le suivi d'objet à partir d'un ou plusieurs capteurs</b>	<b>45</b>
3.1	Modélisation du problème de suivi d'objet . . . . .	45
3.1.1	Méthodes inverses . . . . .	45
3.1.2	Méthodes directes . . . . .	46
3.2	Formalisation du filtre bayésien . . . . .	47
3.3	Méthodes paramétriques . . . . .	48
3.3.1	Modèle linéaire avec bruit gaussien : filtre de Kalman . . . . .	48
3.3.2	Modèle non-linéaire avec bruit gaussien : filtre de Kalman étendu (EKF) . . . . .	50
3.3.3	Modèle non-linéaire avec bruit gaussien : filtre de Kalman sans parfum (UKF) . . . . .	51
3.3.4	Synthèse . . . . .	52
3.4	Méthodes stochastiques . . . . .	53
3.4.1	Le filtre à particules SIR . . . . .	53
3.4.2	L'échantillonneur à particules MCMC . . . . .	54
3.4.3	Le filtre à particules MCMC . . . . .	56
3.4.4	Synthèse . . . . .	57
3.5	Méthode d'association pour le suivi multi-objets . . . . .	58
3.5.1	Fenêtre de validation des mesures . . . . .	59
3.5.2	Méthodes d'association paramétriques . . . . .	60
3.5.3	Méthodes d'association non paramétriques . . . . .	61
3.5.4	Méthodes stochastiques . . . . .	62
3.5.5	Synthèse . . . . .	63
3.6	Extension du problème de l'association de données à la fusion multi-capteurs . . . . .	63
3.6.1	Introduction . . . . .	63
3.6.2	La fusion de données : Buts et approches classiques . . . . .	64
3.6.3	Architectures multi-capteurs classiques dans le cadre d'un filtre bayésien . . . . .	64
3.6.4	Synthèse . . . . .	67
3.7	Conclusion . . . . .	68
<b>4</b>	<b>Algorithmes proposés pour la détection et le suivi de piéton</b>	<b>71</b>
4.1	Détection puis suivi de piéton à partir d'un capteur multi-plans laser . . . . .	72
4.1.1	Segmentation . . . . .	72
4.1.2	Extraction de piétons . . . . .	76
4.1.3	Fusion des détections issues des quatre nappes laser . . . . .	81
4.1.4	Association et suivi multi-objets . . . . .	84
4.2	Fusion multisensorielle laser/vidéo . . . . .	90
4.2.1	Introduction . . . . .	90
4.2.2	Description et positionnement des capteurs sur le véhicule . . . . .	91
4.2.3	Contributions et algorithmes proposés . . . . .	91

4.3	Conclusion . . . . .	97
<b>5</b>	<b>Expérimentations et résultats</b>	<b>101</b>
5.1	Obtention d'une vérité terrain . . . . .	102
5.1.1	Objectifs . . . . .	102
5.1.2	Méthodes existantes . . . . .	102
5.1.3	Obtention manuelle de vérité terrain . . . . .	103
5.1.4	Conclusion . . . . .	105
5.2	Comparaison de détection entre la méthode fondée sur la combinaison des quatre plans et la méthode fondée sur un seul plan . . . . .	106
5.2.1	Critères de performance de détection de piéton . . . . .	106
5.2.2	Résultats expérimentaux de l'algorithme de suivi de piéton . . . . .	112
5.2.3	Conclusion . . . . .	115
5.3	Détection et suivi de piéton avec retour dans l'image laser . . . . .	117
5.3.1	Problématique . . . . .	117
5.3.2	Du modèle classique au modèle avec rétroaction . . . . .	120
5.4	Association de données centralisée laser/vidéo . . . . .	124
5.4.1	Stratégie et conséquences . . . . .	126
5.4.2	Implémentation et séquences d'expérimentations . . . . .	127
5.4.3	Synthèse : avantages, limitations et contraintes . . . . .	131
5.5	Conclusion . . . . .	133
	<b>Conclusion générale et perspectives</b>	<b>135</b>
<b>A</b>	<b>Bilan représentatif des principaux capteurs laser proposés sur le marché</b>	<b>141</b>
<b>B</b>	<b>Modèles d'évolution pour la trajectographie d'entités piéton</b>	<b>143</b>
B.1	Etude du déplacement d'un piéton . . . . .	143
B.1.1	Modèle d'évolution à vitesse constante . . . . .	143
B.1.2	Modèle d'évolution à mouvement circulaire . . . . .	144
B.2	Evaluation des trajectoires et comparaisons . . . . .	145
B.2.1	Simulations de trajectoires . . . . .	145
B.2.2	Résultats . . . . .	146
B.3	Conclusion . . . . .	149
<b>C</b>	<b>Comparaison des différents algorithmes de détection laser livrés par les par- tenaires du projet LOVE</b>	<b>151</b>
C.1	Evaluation quantitative des modules définissant des détections de zones d'intérêts . . . . .	151
C.2	Evaluation quantitative des modules de reconnaissance de piétons . . . . .	152
C.3	Résultats . . . . .	153
C.3.1	Taux de détections de zones d'intérêts . . . . .	153
C.3.2	Taux de reconnaissance des piétons . . . . .	154
	<b>Bibliographie</b>	<b>157</b>



# Table des figures

1.1	Risque de mortalité des piétons en fonction de la vitesse d'impact d'une voiture [Paden <i>et al.</i> , 2004]. . . . .	3
1.2	Exemple de résultats EuroNcap du 18 février 2009 sur les derniers véhicules mis sur le marché - figure extraite de <a href="http://www.euroncap.com">www.euroncap.com</a> . . . . .	5
1.3	LOVe est un projet ANR PREDIT d'une durée de trois ans labelisé par Systém@tic et soutenu par MOV'EO et ViaMéca . . . . .	8
1.4	Exemple d'images laser enregistrée par un capteur laser quatre plans en contexte urbain. Les quatre couleurs permettent de distinguer les quatre plans du LIDAR. . . . .	10
1.5	Exemple d'un suivi de piéton avec deux caméras. Gauche : caméra 1, milieu : caméra 2, droite : trajectoires en vue de dessus - figure extraite de [Bardet, 2009] . . . . .	11
1.6	Vue d'ensemble de la thèse. . . . .	12
2.1	Système de coordonnées cartésiennes et pseudo-sphériques. . . . .	17
2.2	A gauche : un véhicule VIPA du LASMEA équipé d'un capteur Sick. A droite : un plan laser obtenu avec le capteur Sick. . . . .	19
2.3	A gauche : le capteur Sick monté autour d'un axe de rotation vertical. A droite : un résultat obtenu à partir d'un capteur Sick en rotation azimuth - figure extraite de [Wulf & Wagner, 2003]. . . . .	19
2.4	A gauche : le véhicule VELAC. A droite : un résultat obtenu à partir d'un capteur IBEO 4 plans. Chaque couleur correspondant à un plan laser. . . . .	20
2.5	Image d'intensité dense d'un bureau - figure extraite de [Blanc, 2005]. . . . .	21
2.6	A gauche, le capteur IBEO LD ML qui équipe le véhicule VELAC du LASMEA. A droite, le capteur IBEO ALASCA XT qui équipe un des véhicules tests de Renault. . . . .	21
2.7	Capteur laser multi-plans avec quatre plans répartis sur un angle d'ouverture de $3,2^\circ$ . . . . .	21
2.8	Position d'un plan laser en fonction de l'angle de tangage du véhicule. Comme le montre cette figure, un piéton localisé à 30 m ne peut pas être détecté avec un angle de tangage de $3.2^\circ$ . . . . .	22
2.9	Exemple de segmentation d'une image d'impacts laser en vue de dessus. Un groupe de points doit présenter au minimum trois points consécutifs. . . . .	23
2.10	Schéma des deux approches permettant d'estimer l'état $X$ qui représente notre compréhension de la scène. L'estimation de l'état à partir d'une observation capteur est un problème inverse. Le processus qui consiste à construire une observation à partir d'un <i>a priori</i> sur la carte ou sur les objets cherchés est un problème direct. . . . .	24

2.11	Principe d'une mise en correspondance entre deux images laser consécutives. A gauche, deux images laser consécutives sans mise en correspondance A droite, résultat après une mise en correspondance - figure extraite de [Gutmann & Schlegel, 1996]. . . . .	24
2.12	Résultat d'une détection d'arcs (à gauche) et de lignes (à droite) - figure extraite de [Xavier <i>et al.</i> , 2004]. . . . .	25
2.13	A gauche, situation d'une apparition brutale d'un enfant - figure extraite de [Gavrila, 2001] . A droite, passage piéton avec de nombreuses occultations entre les usagers. . . . .	27
2.14	Exemples de différentes postures que peut prendre un piéton - figure extraite de [Gavrila <i>et al.</i> , 2001]. . . . .	27
2.15	Ces deux images illustrent pourquoi la détection puis la classification des objets sont compliquées avec un capteur laser IBEO alasca XT. En effet, l'environnement perçu par ce capteur est « pauvre » c'est à dire que nous n'avons pas d'information sur la couleur, la hauteur ou la forme des objets. . . . .	28
2.16	Nombre de points théoriques (pour une nappe) renvoyés par un piéton de profil (largeur $\sim 40$ cm) pour un angle de résolution égal à $0.5^\circ$ en fonction de la distance séparant le piéton du capteur. . . . .	29
2.17	Le nombre de points renvoyés par un piéton dans une image laser peut être très différent selon sa posture et sa position relative par rapport au capteur. Par exemple, le piéton de gauche étant plus proche du capteur de détection que le piéton de droite, le capteur devrait renvoyer plus de points. Mais comme le piéton est de profil (surface de réflexion réduite), il renvoie moins de points que le piéton de face. . . . .	29
2.18	Synthèse bibliographique des principales méthodes de détection de piéton à partir d'un capteur laser. Ces méthodes sont classées de gauche à droite en fonction des contextes expérimentaux. Notre travail est situé à droite de ce tableau, c'est-à-dire dans le contexte le plus général. . . . .	31
2.19	Image du fauteuil roulant de Prassler circulant dans une station de métro - figure extraite de [Prassler <i>et al.</i> , 1999]. . . . .	32
2.20	Suivi de personnes en intérieur avec à gauche une image de la séquence et à droite une visualisation 3D de la grille d'occupation - figure extraite de [Schulz <i>et al.</i> , 2001]. . . . .	32
2.21	Exemple d'une grille d'occupation utilisée par Lindstöm pour détecter des obstacles en mouvement à partir de la violation des points - figure extraite de [Lindstrom & Eklundh, 2001]. . . . .	33
2.22	Modèle de marche d'un piéton - figure extraite de [Fuerstenberg & Dietmayer, 2004]. . . . .	33
2.23	Détection des jambes d'un piéton dans une image laser à différents instants - figure extraite de [Fuerstenberg & Dietmayer, 2004]. . . . .	33
2.24	Enregistrement du mouvement de la jambe gauche (en rouge) et de la jambe droite (en vert) d'un piéton - figure extraite de [Fuerstenberg & Dietmayer, 2004]. . . . .	34
2.25	Le système de détection de piéton de Shao à partir d'un capteur laser. Dans l'image en bas à droite, nous pouvons remarquer le mouvement périodique des jambes qui a été détecté puis suivi - figures extraites de [Shao <i>et al.</i> , 2007] . . . . .	34

2.26	Diverses situations considérées comme dangereuses par les travaux de Broggi. Par exemple en (a), au niveau d'un passage piéton, un piéton est caché par un véhicule garé ; en (b), un autre est en train de traverser la route derrière un arrêt de bus ; en (c) et (d), un troisième traverse la route en passant entre deux véhicules garés au bord de la route - figures extraites de [Broggi <i>et al.</i> , 2008]. . . . .	35
2.27	Illustration de la segmentation en ligne de l'image laser (à gauche), et classification (à droite) dans des boîtes englobantes des objets détectés dans l'image à partir du calcul du seuil - figures extraites de [Navarro-Serment <i>et al.</i> , 2008]. . . . .	35
2.28	Fonction permettant de calculer la note de reconnaissance en fonction de la largeur du piéton - figure extraite de [Fayad <i>et al.</i> , 2008]. . . . .	36
2.29	A gauche, exemple d'une détection et d'un suivi de piéton en vue de dessus dans un environnement intérieur. A droite, résultats temporels du suivi des objets dans cette pièce. Les étoiles représentent la vérité terrain tandis que les lignes représentent le résultat de détection et de suivi de l'algorithme - figures extraites de [Fod <i>et al.</i> , 2002]. . . . .	36
2.30	Illustration à gauche d'une détection de piéton en comparant un groupe de points avec un cylindre. A droite, exemple d'une détection et du suivi de deux piétons à partir du robot mobile - figures extraites de [Montemerlo <i>et al.</i> , 2002]. . . . .	37
2.31	Modèle de la fonction de « Vote à la majorité » utilisée par [Mendes <i>et al.</i> , 2004]. . . . .	37
2.32	Résultats expérimentaux de SLAM avec des objets en mouvements - figures extraites de [Zhao <i>et al.</i> , 2008]. . . . .	38
2.33	Détection de piéton à partir d'un capteur laser et de la méthode SLAM-MOT. A gauche, une photo du lieu observé et à droite, les détections de piétons effectuées dans l'image laser - figures extraites de [Gate & Nashashibi, 2008]. . . . .	39
2.34	Illustration d'une base de données comportant plusieurs « imageries » de la trace d'un piéton dans une image d'impacts laser en vue de dessus utilisée par le classifieur Adaboost - figure extraite de [Spinello <i>et al.</i> , 2008]. . . . .	39
2.35	En noir, schéma d'un système de détection de piéton « idéal ». Un résultat probable d'algorithme de détection est représenté en vert avec son lot de fausses détections (contour en dehors du rectangle noir) et de non détection (contour à l'intérieur du rectangle noir). En pointillé, est représenté un système parfait de détection de piéton situé en situation de collision avec le véhicule alors qu'un résultat probable d'algorithme de détection est représenté en rouge avec également son lot de fausses détections et de non détection. . . . .	41
2.36	Exemple typique d'une scène observée dans une caméra opérant dans le visible (à gauche) et dans l'infrarouge (à droite) - figures extraites de [Bertozzi <i>et al.</i> , 2007] . . . . .	42
2.37	En vert, un piéton qui doit être absolument détecté (buste visible) et en rouge, un piéton masqué dont la détection n'est pas requise <i>a priori</i> . . . .	42

2.38	Un piéton est à risque si et seulement s'il est présent dans la zone rouge. Cette zone est définie de la façon suivante : on considère un piéton se déplaçant à $2\text{ m/s}$ perpendiculairement à la trajectoire du véhicule. La dimension du pare-chocs étant constante (par exemple $L = 2\text{ m}$ ), les piétons pouvant être heurtés sont à une distance $V_{\text{piéton}} * t_{\text{avant\_impact}}$ du pare-chocs soit $2 * 2 = 4$ mètres. . . . .	43
3.1	Schéma des deux approches permettant d'estimer l'état de la variable cachée qui représente notre connaissance de la scène. L'estimation de l'état à partir d'une observation est un problème inverse. Le processus qui consiste à construire une observation à partir de l'état est un problème direct. . . . .	46
3.2	Modèle graphique du réseau bayésien dynamique pour un suivi multi-objets.	47
3.3	Représentation schématique du filtre bayésien récursif par les méthodes paramétriques. . . . .	48
3.4	Représentation schématique du filtre bayésien récursif par les méthodes séquentielles de Monte Carlo. . . . .	49
3.5	Propagation des particules par l'algorithme SIR - figure tirée de [Smith, 2007]. . . . .	54
3.6	Approximation d'une distribution cible $p(X)$ par un échantillonneur MCMC. En (a), la distribution cible $p(X)$ est présentée. En (b), le résultat de l'approximation MCMC par Metropolis-Hasting de la distribution cible pour un nombre suffisamment grand d'échantillons $N$ et après élimination des $N_B$ premiers échantillons ( <i>burn-in</i> ) . La distribution est bien approximée. En (c), résultat de l'approximation MCMC par Metropolis-Hasting de la distribution cible mais avec un nombre insuffisant d'échantillons $N$ . La distribution est alors mal approximée. Enfin, en (d), résultat de l'approximation MCMC par Metropolis-Hasting de la distribution cible pour un nombre suffisamment grand d'échantillons $N$ mais sans élimination des $N_B$ premiers échantillons. La distribution se retrouve alors biaisée par le point de départ (l'initialisation) de notre chaîne - figure extraite de [Smith, 2007]. . . . .	56
3.7	Le module de suivi. . . . .	58
3.8	L'Association par PDAF consiste à construire un estimateur qui prend en compte toutes les observations de la fenêtre, donc toute l'information contenue dans celle-ci. Sur cette figure, on peut voir que la piste est associée avec une donnée « fictive » calculée en pondérant chaque donnée présente dans la fenêtre de validation. . . . .	60
3.9	Chevauchement de deux fenêtres de validation pour deux objets et cinq mesures. . . . .	61
3.10	Exemple d'une association par PPV. . . . .	62
3.11	Exemple d'associations par MHT. . . . .	63
3.12	Architecture de fusion série entre un capteur laser et un capteur vidéo. Les observations laser alimentent la recherche vidéo avant la phase de poursuite multi-pistes. . . . .	65
3.13	Architecture de fusion centralisée entre un capteur laser et un capteur vidéo. Les observations des deux capteurs sont traitées dans le module d'association "globale" avant la phase de poursuite multi-pistes. . . . .	66



3.14	Architecture de fusion décentralisée entre un capteur laser et un capteur vidéo. Chaque capteur fait sa propre poursuite multi-objets, les pistes des deux capteurs sont ensuite fusionnées. . . . .	67
3.15	Architecture de fusion hybride entre un capteur laser et un capteur vidéo. Chaque capteur fait sa propre poursuite multi-objets. Les observations et les pistes des deux capteurs sont traitées dans le module d'association "globale" avant la phase de poursuite multi-pistes. . . . .	68
4.1	Algorithme de détection puis de suivi de piéton à partir d'un capteur multi-plans laser. La figure 4.2 illustre ces différentes étapes. . . . .	72
4.2	Illustration des différentes étapes de l'algorithme de détection puis de suivi de piéton à partir d'un capteur multi-plans laser. En bas, les piétons détectés qui sont reprojétés dans l'image vidéo. . . . .	73
4.3	Résultats d'une extraction de segments dans une image laser. . . . .	75
4.4	Résultats d'une extraction d'amas de points dans une image laser. . . . .	76
4.5	Exemple d'une image laser comportant plusieurs piétons à détecter. Dans une image de profondeur un « objet-piéton » est représenté par un amas de points répartis géométriquement en fonction de sa largeur, de sa distance avec le capteur et de sa posture. Sans vérité terrain, un homme risque de détecter tous les carrés bleus. . . . .	78
4.6	Exemple d'une extraction d'« objets-piétons » à partir de l'image laser ci-dessus, obtenue à partir d'une densité non paramétrique à noyaux comportant les caractéristiques géométriques d'un piéton (cf. l'équation 4.19). . . . .	78
4.7	Les quatre images laser représentent la scène observée suivant les quatre plans laser du capteur. Les piétons détectés par le module d'extraction de piéton sont représentés par une étoile noire. Le résultat final de la fusion des quatre nappes est montré en bas à gauche par un cercle noir. Finalement, le résultat final de détection est projeté dans l'image vidéo qui permet de vérifier l'exactitude du résultat. Il est intéressant de remarquer la différence des informations contenues dans les quatre plans laser de la même scène observée. Le nombre de piétons détectés et leurs positions sont donc différents d'un plan à l'autre. La fusion des quatre plans permet (dans cet exemple), d'éliminer les fausses détections et de ne conserver que les trois piétons. . . . .	82
4.8	Nombre de nappes $N_l$ en fonction de la distance objet/capteur. . . . .	83
4.9	Algorithme de suivi de piéton. . . . .	85
4.10	Exemple d'une matrice d'associations. . . . .	87



4.11	Exemple d'évolution d'une matrice d'associations $A_k$ au temps $k$ . Dans le cas où plusieurs associations sont possible (une colonne de $A_k$ contient plusieurs $d_{ij}$ ), on commence par associer le couple mesure-objet qui possède la plus forte probabilité d'association $d_{ij}$ . Dans le cas de notre espace joint (hypothèses <b>H1</b> et <b>H2</b> ), la distance $d_{ij}$ possédant la plus forte probabilité est $d_{14}$ . Le couple associant l'objet numéro 1 avec l'observation numéro 4 est alors éliminé avec sa ligne et sa colonne complète afin de supprimer toute correspondance future d'une mesure avec d'autres objets. L'algorithme est réitéré tant qu'une observation reste à associer. On remarque que dans notre exemple l'observation numéro 1 ne trouve finalement pas d'objet avec qui s'associer. Dans ce cas, le module de gestion des pistes décide d'en créer une nouvelle. . . . .	89
4.12	Architecture de fusion centralisée fondée sur un filtre particulaire SIR et une association non paramétrique des données. L'algorithme fusionne des mesures provenant d'algorithmes de détection et/ou de reconnaissance laser et vidéo. . . . .	93
4.13	Description de la synchronisation des données d'un capteur sur l'autre dans une échelle commune. . . . .	94
4.14	Description de la non synchronisation des données d'un capteur sur l'autre dans une échelle commune. . . . .	94
4.15	Un exemple d'un mélange de distributions gaussienne et uniforme, avec en bleu la distribution uniforme, en vert la distribution gaussienne et en rouge la distribution mélangée. A gauche, $\sigma = 0.15$ m et $NDG = 0.95$ . A droite, $\sigma = 0.15$ m et $NDG = 0.55$ . . . . .	96
5.1	Exemples de scénarios produits par le simulateur SiVIC - figure extraite de <a href="http://www.inrets.fr/ur/livic/">www.inrets.fr/ur/livic/</a> . . . . .	103
5.2	Extraits d'images vidéos et laser associées au scénario <b>VIPA campus d'Aubière</b> . Seuls les piétons encadrés en jaune font partie de la vérité terrain. A gauche : exemple d'une détection de trois piétons qui sont tous visibles au sens des spécifications LOVE. A droite : exemple d'un piéton encadré en rouge, visible ni dans la vidéo ni dans l'image de profondeur. Ce piéton n'appartient donc pas à la vérité terrain. . . . .	104
5.3	Extraits d'images vidéos et laser associées au scénario <b>Renault centre-ville</b> . Seuls les piétons encadrés en jaune font partie de la vérité terrain. A gauche : exemple d'une détection d'un piéton visible au sens des spécifications LOVE. A droite : exemple d'un piéton encadré en rouge, visible dans l'image de profondeur mais pas dans la vidéo. Ce piéton n'appartient donc pas à la vérité terrain. . . . .	105
5.4	Illustration de la vérité terrain. . . . .	107
5.5	Relevés vidéo et laser pour l'image #1 du scénario <b>VIPA campus d'Aubière</b> . Ici, une fausse détection. . . . .	108
5.6	Relevés vidéo et laser pour l'image #160 du scénario <b>VIPA campus d'Aubière</b> . Ici, trois détections de piéton confirmées par la vérité terrain et aucune fausse détection. . . . .	108
5.7	Relevés vidéo et laser pour l'image #179 du scénario <b>VIPA campus d'Aubière</b> . Ici, trois détections de piéton confirmées par la vérité terrain et aucune fausse détection. . . . .	109

5.8	Relevés vidéo et laser pour l'image #218 du scénario <b>VIPA campus d'Aubière</b> . Ici, deux détections de piéton confirmées par la vérité terrain.	109
5.9	Relevés vidéo et laser pour l'image #228 du scénario <b>VIPA campus d'Aubière</b> . Ici, deux détections de piéton confirmées par la vérité terrain et une fausse détection.	109
5.10	Relevés vidéo et laser pour l'image #268 du scénario <b>VIPA campus d'Aubière</b> . Ici, deux détections de piéton confirmées par la vérité terrain et une fausse détection.	110
5.11	Relevés vidéo et laser pour l'image #1 du scénario <b>Renault centre-ville</b> . Ici, aucune détection de piéton et aucune fausse détection.	110
5.12	Relevés vidéo et laser pour l'image #11 du scénario <b>Renault centre-ville</b> . Ici, une détection de piéton confirmée par la vérité terrain.	110
5.13	Relevés vidéo et laser pour l'image #16 du scénario <b>Renault centre-ville</b> . Ici, aucune détection de piéton et aucune fausse détection.	111
5.14	Relevés vidéo et laser pour l'image #22 du scénario <b>Renault centre-ville</b> . Ici, deux détections de piéton et une fausse détection.	111
5.15	Relevés vidéo et laser pour l'image #32 du scénario <b>Renault centre-ville</b> . Ici, une détection de piéton et aucune fausse détection.	111
5.16	Le premier scénario présente le suivi de trois piétons situés dans un environnement semi-urbain en utilisant l'algorithme de détection fondé sur les quatre nappes laser. Ce scénario est enregistré à bord du véhicule expérimental du LASMEA. Les principales difficultés de ce scénario sont le filtrage des arbustes et les fréquentes occultations des piétons entre eux. Les piétons détectés dans l'image laser sont cerclés de noir. Les rectangles bleus représentent leur projection dans l'image vidéo.	114
5.17	Le second scénario présente le suivi de plusieurs piétons situés dans un environnement urbain en utilisant l'algorithme de détection fondé sur les quatre nappes laser. Ce scénario a été enregistré à bord du véhicule expérimental de Renault dans des conditions de circulation à vitesse réelle. Les principales difficultés rencontrées sont le filtrage des gros poteaux de circulation, les véhicules garés sur les bas-côtés, les fréquentes occultations des piétons entre eux et de manière générale, la richesse de l'environnement urbain qui induit de nombreuses détections assimilables à un piéton. Nous reverrons ce scénario plus en détails lors de l'évaluation de l'algorithme de fusion centralisée laser/vidéo. Les piétons détectés dans l'image laser sont cerclés de noir. Les rectangles blancs représentent leur projection dans l'image vidéo. On rappelle que les points de couleur présents dans l'image vidéo représentent la projection des détections effectuées dans les quatre plans.	115
5.18	Résultat du suivi d'un piéton avec l'algorithme SIR/KDE : vue de dessus. Les mesures sont représentées par des cercles gris.	116
5.19	Résultats en positions ( $\rho$ et $\theta$ ) du suivi d'un piéton. Les mesures sont représentées par des cercles gris.	116
5.20	Résultats de l'erreur en positions ( $\rho$ et $\theta$ ) du suivi d'un piéton.	117
5.21	Résultats de la covariance en positions ( $\rho$ et $\theta$ ) du suivi d'un piéton.	117
5.22	Relevés vidéo et laser pour l'image #1.	117
5.23	Relevés vidéo et laser pour l'image #5.	118

5.24	Relevés vidéo et laser pour l'image #35. . . . .	118
5.25	Relevés vidéo et laser pour l'image #47. . . . .	118
5.26	Relevés vidéo et laser pour l'image #113. . . . .	119
5.27	Relevés vidéo et laser pour l'image #139. . . . .	119
5.28	Relevés vidéo et laser pour l'image #157. . . . .	119
5.29	Relevés vidéo et laser pour l'image #234. . . . .	119
5.30	Relevés vidéo et laser pour l'image #278. . . . .	120
5.31	Relevés vidéo et laser pour l'image #296. . . . .	120
5.32	Relevés vidéo et laser pour l'image #345. . . . .	120
5.33	Schéma de l'algorithme de détection et de suivi de piéton avec <i>feedback</i> dans l'image laser. . . . .	121
5.34	Les piétons détectés sont représentés par des points noirs tandis que les piétons suivis sont représentés par des rectangles de couleur. La couleur permet d'identifier la piste. A gauche, extrait d'un suivi de piéton dans une image laser avec l'algorithme sans boucle de retour. Ici, quand le piéton en second plan est occulté, le module de pistage sans <i>feedback</i> perd sa trace car il n'est pas détecté par le module de détection. Lorsque ce piéton redevient observable par le module de détection, le module de suivi propose d'initialiser une nouvelle piste, munie d'une nouvelle identité (rectangle qui passe du rouge au bleu). A droite, même extrait avec l'algorithme utilisant la boucle de retour. Cette fois, le module de pistage avec <i>feedback</i> arrive à garder une trace du piéton en le cherchant dans l'image laser permettant d'obtenir une mesure capable de s'associer avec la piste qui survit et conserve son identité (pas de changement de couleur du rectangle). . . . .	122
5.35	Superposition des 2 courbes ROC permettant de comparer les deux algorithmes de suivi (avec et sans <i>feedback</i> ) sur le scénario <b>VIPA campus d'Aubière</b> . . . . .	124
5.36	Résultats du suivi de deux piétons sur le scénario commenté dans la section 5.2.2 avec l'algorithme sans boucle de retour. Les mesures obtenues par l'algorithme de détection quatre plans sont représentées en gris. Les pistes sont identifiées par des étoiles de couleur. . . . .	125
5.37	Résultats du suivi de deux piétons sur le scénario commenté dans la section 5.2.2 avec l'algorithme utilisant une boucle de retour. Les mesures obtenues par l'algorithme de détection quatre plans sont représentées en gris. Les pistes sont identifiées par des étoiles de couleur. . . . .	125
5.38	Résultats temporels des deux algorithmes de suivi (avec et sans <i>feedback</i> ) sur le scénario <b>VIPA campus d'Aubière</b> . . . . .	126
5.39	Séquence 1 : Résultats de l'algorithme de fusion centralisée sur un scénario livré par Renault. . . . .	128
5.40	Séquence 1 : résultats du pistage en x et y : cas mono-cible d'une fusion centralisée laser-vidéo. Les mesures (vidéo et laser) sont représentées par des cercles gris. Chaque piste est représentée par une couleur différente. . . . .	129
5.41	Séquence 1 : résultats du Facteur de Confiance sur la Fusion de Données (FCFD) associé à la figure 5.43. Chaque FCFD est représenté par une étoile de la même couleur que sa piste (cf. la figure 5.43). . . . .	129

5.42	Séquence 2 : résultats de l'algorithme de fusion centralisée sur un scénario livré par Renault. . . . .	131
5.43	Séquence 2 : résultats du pistage en x et y : cas mono-cible d'une fusion centralisée laser-vidéo. Les mesures (vidéo et laser) sont représentées par des cercles gris. Chaque piste est représentée par une couleur différente. .	132
5.44	Séquence 2 : résultats du Facteur de Confiance sur la Fusion de Données (FCFD) associé à la figure 5.43. Chaque FCFD est représenté par une étoile de la même couleur que sa piste (cf. la figure 5.43). . . . .	132
B.1	Scénario qui propose le suivi d'un piéton sur la trajectoire en forme de courbe. A gauche, vue de dessus du résultat de suivi en position. A droite, vue de dessus du résultat de suivi en vitesse. Les mesures sont représentées par des cercles gris. . . . .	146
B.2	Résultats de l'algorithme de suivi SIR/modèle à vitesse constante de la figure B.1. A gauche en haut : l'erreur en position. A gauche en bas : l'erreur en vitesse. A droite l'efficacité de l'échantillonnage. . . . .	147
B.3	Scénario qui propose le suivi d'un piéton sur une trajectoire complexe comportant un changement brusque de direction. A gauche, résultat du suivi en position et à droite, résultat du suivi en vitesse. Les mesures sont représentées par des cercles gris. . . . .	147
B.4	Résultats de l'algorithme de suivi SIR/modèle à vitesse constante de la figure B.3. A gauche, l'erreur en position et en vitesse, à droite l'efficacité de l'échantillonnage. . . . .	147
B.5	Scénario qui propose le suivi d'un piéton sur la trajectoire en forme de courbe. A gauche, vue de dessus du résultat de suivi en position. A droite, vue de dessus du résultat de suivi en vitesse. Les mesures sont représentées par des cercles gris. . . . .	148
B.6	Résultats de l'algorithme de suivi SIR/modèle à vitesse constante de la figure B.5. A gauche en haut : l'erreur en position. A gauche en bas : l'erreur en vitesse. A droite l'efficacité de l'échantillonnage. . . . .	148
B.7	Scénario qui propose le suivi d'un piéton sur une trajectoire complexe comportant un changement brusque de direction. A gauche, résultat du suivi en position et à droite, résultat du suivi en vitesse. Les mesures sont représentées par des cercles gris. . . . .	149
B.8	Résultats de l'algorithme de suivi SIR/modèle à vitesse constante de la figure B.7. A gauche, l'erreur en position et en vitesse, à droite l'efficacité de l'échantillonnage. . . . .	149
C.1	Courbe de caractérisation des performances de détection des zones d'intérêts. . . . .	152
C.2	Courbe ROC de caractérisation des performances en reconnaissance. . . .	153
C.3	Taux de sélectivité en fonction du nombre de zones d'intérêts détectées sur quatre scénarios Renault. Le nombre de zone d'intérêts « vérité terrain » est précisé par la variable « Nb GT ». Le nombre total de détections est donné par la variable « Nb algo ». . . . .	154
C.4	Courbes ROC correspondant à la reconnaissance des piétons. . . . .	155



---

# Liste des tableaux

1.1	Nombres de piétons tués et blessés en France depuis vingt ans, <a href="http://www.securite-routiere.gouv.fr">www.securite-routiere.gouv.fr</a> . . . . .	2
2.1	Etat de l’art des capteurs disponible au LASMEA et/ou dans le contexte du projet LOVE. . . . .	18
3.1	Principaux avantages et inconvénients des quatres architectures de fusion entre un laser et une caméra. . . . .	68
5.1	Taux de fausses détections et de bonnes détections en fonction du nombre de nappes utilisées sur le scénario <b>VIPA campus d’Aubière</b> . . . . .	108
5.2	Taux de fausses détections et de bonnes détections en fonction du nombre de nappes utilisées sur le scénario <b>Renault centre-ville</b> . . . . .	112
5.3	Séquence 1 : taux de fausses et de bonnes détections sur scénario livré par Renault. . . . .	130
5.4	Séquence 2 : taux de fausses et de bonnes détections sur un scénario livré par Renault. . . . .	131



---

# Liste des algorithmes

1	Filtre de Kalman . . . . .	50
2	Filtre de Kalman étendu . . . . .	51
3	Filtre de Kalman sans parfum . . . . .	52
4	Filtre particulaire SIR (Sequential Importance Resampling) . . . . .	54
5	L’algorithme de Metropolis-Hasting pour un échantillonneur MCMC . . .	55
6	L’algorithme du filtre MCMC . . . . .	57
7	Algorithme de détection de piétons avec un estimateur non paramétrique à noyau . . . . .	80
8	Algorithme de détection d’« objets-piétons » issues des quatre nappes laser avec un estimateur non paramétrique à noyau . . . . .	84
9	Association et suivi temporel : SIR / KDE. . . . .	90
10	Fusion centralisée non paramétrique des données. . . . .	98
11	Module de détection avec <i>feedback</i> dans l’image laser. . . . .	123





---

# Notations et Conventions

Tout au long de ce manuscrit, les vecteurs sont représentés par des lettres en gras alors que les variables sont en fonte normale. Les acronymes suivants sont précédés de leur équivalent français lorsqu'ils sont issus de la langue anglaise :

## Acronymes

BDDF ....	Fusion de données décentralisée bayésienne, Bayesian Decentralized Data Fusion.
EKF .....	Filtre de Kalman étendu, Extended Kalman Filter.
EM .....	Algorithme d'espérance-maximisation, Expectation-Maximisation algorithm.
PF .....	Filtre particulaire, Particle Filter.
GMM ....	Modèle de mélanges gaussiens, Gaussian Mixture Model.
JPDAF ...	Filtre probabiliste joint à association de données, Joint Probabilistic Data Association Filter.
KDE .....	Estimation de densité par la méthode du noyau, Kernel Density Estimation.
KF .....	Filtre de Kalman, Kalman Filter.
Lidar .....	Capteur de distance à balayage laser, Light Detection and Ranging.
LOVe ....	Logiciel d'Observation des Vulnérables.
MAP .....	Maximum a posteriori, Maximum A Posteriori.
MCMC ...	Méthodes de Monte-Carlo par chaînes de Markov, Markov Chain Monte Carlo.
MHT .....	Filtre à hypothèses multiples, Multiple Hypotheses Tracking.
MLE .....	Estimateur par maximum de vraisemblance, Maximum Likelihood Estimator.
MOT .....	Suivi d'objet en mouvement, Moving Object Tracking.
MSE .....	Erreur quadratique moyenne, Mean Squared Error.
MTT .....	Suivi d'objets multiples, Multiple Target Tracking.
PDAF ....	Filtre probabiliste à association de données, Probabilistic Data Association Filter.
PMHT ...	Filtre probabiliste à hypothèses multiples, Probabilistic Multiple Hypotheses Tracking.
PPV .....	Méthode du Plus Proche Voisin.
SIR .....	Echantillonnage pondéré séquentiel avec rééchantillonnage, Sequential Importance Resampling.
SIS .....	Echantillonnage pondéré séquentiel, Sequential Importance Sampling.
UKF .....	Filtre de Kalman "unscented", Unscented Kalman Filter.
VIPA .....	Vehicules Individuels Publics Automatiques.

## Notations

$\mathbb{N}$ .....	Ensemble des entiers naturels.
$\mathbb{R}$ .....	Ensemble des nombres réels.
$(\Omega, \Theta)$ .....	Espace probabilisé sur lequel sont définies les variables aléatoires utilisées.
$\mathbb{E}_p$ .....	Espérance calculée par rapport à la loi de densité $p$ .
$k$ .....	Lettre utilisée en indice pour désigner le temps discret.
$i$ .....	Lettre utilisée en exposant pour désigner un des $M$ objets.
$j$ .....	Lettre utilisée en exposant pour désigner une des $m_t$ mesures à l'instant $t$ .
$n$ .....	Lettre utilisée en exposant pour désigner une des $N$ particules.
$p$ .....	Lettre utilisée en exposant pour désigner une des $P$ réalisations d'une méthode particulière.
$\mathbf{X}_k$ .....	Vecteur d'état Markovien à estimer.
$\mathbf{F}_k$ .....	Fonction d'évolution (linéaire ou non-linéaire) du vecteur d'état.
$\mathbf{V}_k$ .....	Bruit d'état.
$\Sigma$ .....	Matrice de covariance.
$\mathbf{Z}_k$ .....	Vecteur de mesure.
$\mathbf{H}_k$ .....	Fonction de mesure reliant une mesure $z_k^j$ à l'état dont elle dépend.
$\mathbf{W}_k$ .....	Bruit de mesure.
$\mathbf{R}_k$ .....	Matrice de covariance du bruit de mesure.
$\mathbf{Q}_k$ .....	Matrice de covariance du bruit de process.
$\mathbf{S}_k$ .....	Matrice de covariance de l'innovation.
$w_k$ .....	Variable aléatoire de Bernouilli.
$\Phi$ .....	Fonction noyau.
$\mathbf{x}_k^i$ .....	Etat de la cible $i$ à l'instant $k$ .
$\mathbf{x}_k^{i,n}$ .....	Particule $n$ de la cible $i$ à l'instant $k$ .
$\tilde{\mathbf{x}}_k^{i,n}$ .....	Particule prédite à partir de $\mathbf{x}_k^{i,n}$ .
$\tilde{\pi}_k^{i,n}$ .....	Poids non normalisé de la particule $\mathbf{x}_k^{i,n}$ .
$\pi_k^{i,n}$ .....	Poids normalisé de la particule $\mathbf{x}_k^{i,n}$ .
$p(\mathbf{X}_0)$ .....	Etat initial d'une distribution de probabilité.
$\mathcal{N}(\mu, \Sigma)$ .....	Distribution Gaussienne de moyenne $\mu$ et de covariance $\Sigma$ .
$\mathcal{U}(a, b)$ .....	Distribution uniforme sur l'intervalle $[a, b]$ .
$\delta$ .....	Fonction dirac.
$\mathbf{A}_k^t$ .....	Transposé de la matrice $\mathbf{A}_k$ .
$\mathbf{A}_k^{-1}$ .....	Inverse de la matrice $\mathbf{A}_k$ .
$p(\mathbf{Z}_k   \mathbf{X}_k)$ .....	Fonction de vraisemblance.
$p(\mathbf{X}_k   \mathbf{Z}_{1:k})$ .....	Distribution de densité <i>a posteriori</i> .
$p(\mathbf{X}_k   \mathbf{X}_{k-1})$ .....	Fonction d'évolution.
$\arg \max_x (p(\mathbf{Z}_k   \mathbf{X}_k))$	Argument du maximum de vraisemblance.
$\arg \max_x (p(\mathbf{X}_k   \mathbf{Z}_{1:k}))$	Argument du maximum de vraisemblance <i>a posteriori</i> .



---

# Chapitre 1

## Introduction générale

Ce chapitre commence par fournir quelques données d'accidentologie routière sur la part des accidents impliquant un vulnérable (piéton, cycliste, motocycliste) et un véhicule. Nous dressons ensuite un panorama des systèmes de sécurité qui ont investi les véhicules de séries de ces dernières années car elles prennent désormais plus précisément en compte la protection des piétons. Nous exposons ensuite les principaux projets qui concernent la détection des piétons depuis 2002 et notamment le projet ANR <sup>1</sup> LOVe <sup>2</sup> qui constitue le cadre applicatif de cette thèse. Ce chapitre se termine par la présentation de nos travaux et plus précisément des contributions scientifiques des algorithmes que nous avons développés.

### 1.1 Généralités sur la sécurité routière des vulnérables

#### 1.1.1 Les usagers de la route : la perspective européenne

Les routes en mauvais état et des densités de population élevées sont des facteurs souvent cités comme contribuant au taux de mortalité chez les vulnérables de la route, tant dans les pays sous-développés que dans les pays développés. Ces facteurs ne constituent pas réellement un problème en Europe. En tant que zone développée, l'Europe est l'une des zones où il y a le plus de propriétaires de véhicules au monde. Il n'est donc pas surprenant que presque 80% des victimes de la route soient des occupants de véhicules automobiles.

Les vulnérables représentent donc l'autre partie (20%) d'usagers tués ou grièvement blessés chaque année dans des collisions. En 2008, 6000 vulnérables ont été tués et 150 000 ont été blessés. Les piétons représentent quant à eux, 61% des victimes mortellement blessées et 52% de celles blessées grièvement. Les gens qui voyagent à cyclomoteur représentent 28% des décès et 33% des blessés graves. Enfin, on recense 11% de blessures mortelles et 15% de graves chez les cyclistes.

Face à ce triste bilan, la Commission Européenne (CE) a décidé de réagir afin d'améliorer le niveau de sécurité des usagers de la route, considérés comme vulnérables, et s'est fixée en particulier comme objectif d'ici 2010 de réduire le nombre de victimes de la route de 50%. En effet, dans un livre blanc sur les transports paru en 2001, cette même commission a indiqué un certain nombre d'axes de recherche pour améliorer la situation, et

---

<sup>1</sup> Agence Nationale de la Recherche.

<sup>2</sup> Logiciel d'Observation des Vulnérables.

parmi ceux-là notamment les systèmes d'aide à la conduite. Une de ses études montre que les conditions de la protection des piétons peuvent être sensiblement améliorées par le recours à une combinaison de mesures de sécurité actives et passives. Les mesures passives contribuent à réduire la gravité des lésions en cas d'impact par l'utilisation de surfaces plus douces. Les mesures actives atténuent les conditions dans lesquelles l'impact peut se produire, par exemple par la réduction de la vitesse du véhicule. Une telle combinaison de mesures devrait permettre d'améliorer de 80% la protection actuelle des piétons.

Le premier volet de cette réglementation (sécurité passive) est entré en vigueur pour les véhicules neufs le 1er octobre 2005. Il impose de nouvelles normes aux constructeurs automobiles afin de réduire le nombre d'accidents mortels mais surtout de baisser significativement la gravité de ceux qui ne le sont pas. Le second volet de cette réglementation (sécurité active) devrait être fixé en septembre 2017. Il devrait imposer l'intégration de système d'assistance au freinage en cas de collision immédiate avec un piéton.

### 1.1.2 Les principaux usagers concernés

Depuis plus de vingt ans, en France comme dans la majorité des autres pays européens, le nombre de piétons tués dans un accident de la circulation tend à décroître (cf. le tableau 1.1) mais l'enjeu reste important : 580 tués en France en 2008 et un peu plus de 9528 blessés.

Année	Piétons tués	Piétons blessés
1990	1407	19298
1996	987	15020
2000	793	14866
2005	650	13600
2008	580	9528

TAB. 1.1 – Nombres de piétons tués et blessés en France depuis vingt ans, [www.securite-routiere.gouv.fr](http://www.securite-routiere.gouv.fr).

Afin de mieux comprendre quelles sont les principales causes d'accidents de la route impliquant des piétons, l'INRETS <sup>3</sup> [Fontaine, 1995] et le WHO <sup>4</sup> [Paden *et al.*, 2004] ont réalisé une étude qui a permis de mettre en évidence les principaux facteurs de risque d'accident à partir d'une analyse de données fournies par les rapports de police relatant des collisions entre des piétons et des véhicules. De cette étude, les trois points suivant ont attiré notre attention :

- ▷ **Régions urbaines** : près de 70% des piétons sont tués et plus de 90% sont grièvement blessés dans des régions urbaines où les limites de vitesse affichées sont de 70km/h ou moins [Gavrila *et al.*, 2001]. Toujours en région urbaine, le nombre de piétons tués aux intersections est légèrement supérieur à celui de ceux tués ailleurs.
- ▷ **Perceptibilité et période de la journée** : plus de la moitié ont perdu la vie et 40% ont été grièvement blessés la nuit ou dans des conditions d'éclairage artificiel. Environ 20% des piétons victimes de collisions de la route ont été tués entre 17h et 19h.

<sup>3</sup>Institut National de Recherche sur les Transports et leur Sécurité

<sup>4</sup>World Health Organisation

- ▷ **Vitesse d'impact** : des études montrent que plus la vitesse d'impact augmente, plus le risque de lésions graves et mortelles augmente (cf. la figure 1.1). Notamment les piétons ont plus de 90% de chance de survivre à un accident avec une voiture roulant à 30 km/h maximum, mais moins de 50% de chance de survie à l'impact à 45 km/h ou plus. Les piétons âgés sont encore plus vulnérables physiquement aux collisions avec un véhicule. Une vitesse inappropriée contribue à environ 30% des accidents mortels dans les pays développés. En effet, la probabilité qu'un piéton soit tué augmente d'un facteur huit lorsque la vitesse d'impact avec un véhicule augmente de 30 km/h à 50 km/h.

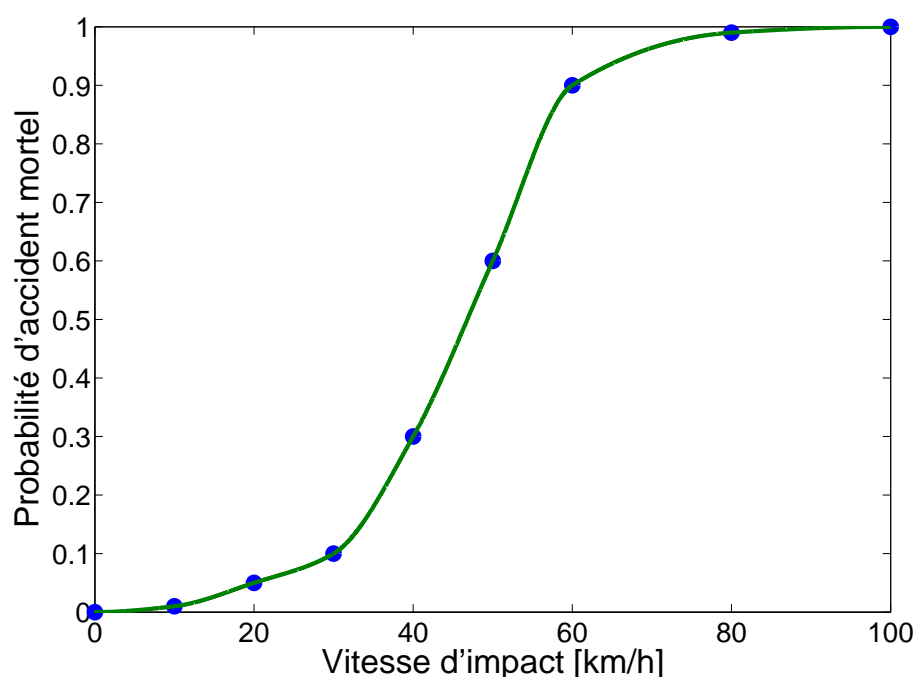


FIG. 1.1 – Risque de mortalité des piétons en fonction de la vitesse d'impact d'une voiture [Paden *et al.*, 2004].

A la lecture de cette étude, il semblerait que l'utilisation d'un système actif de détection de piéton sur un véhicule routier en environnement urbain permettrait une diminution significative du nombre de tués ou, tout au moins, de blessures graves causées par cet accident. En effet, un tel système permettrait un freinage d'urgence du véhicule en cas de situation immédiate d'accident avec un piéton ce qui réduirait considérablement la distance d'arrêt du véhicule. Ainsi la collision pourrait être évitée ou se produire à une vitesse nettement moins élevée.

## 1.2 La protection des piétons dans les véhicules routiers

### 1.2.1 Introduction

Pendant longtemps, la priorité s'est portée sur les systèmes de sécurité passive, qui permettent de limiter les conséquences physiques des accidents. L'airbag et les ceintures



de sécurité sont des exemples qui ont donné satisfaction pour diminuer la gravité des blessures des usagers du véhicules en cas d'accident sur la route. En effet, ces équipements devenus incontournables sont maintenant mis en série sur les véhicules neufs. Ces dernières années, en plus d'améliorer la protection des passagers, les constructeurs cherchent à améliorer la sécurité des vulnérables toujours avec l'aide des systèmes de sécurité passive. Différentes pistes ont été suivies comme celle qui consiste à ajouter un volume supplémentaire entre le capot et les organes mécaniques afin de glisser des absorbeurs de choc en mousse. Le choix des matériaux doit alors être aussi résistants qu'élastiques et faciles à réparer. Une autre approche consiste à aménager un capot qui se relève automatiquement en cas de choc, de façon à éviter au piéton de venir s'encastrent dans le pare-brise. Malheureusement en cas de collision véhicule-piéton, toutes ces améliorations apportées par les systèmes de sécurité passive ne suffisent pas. En effet, en cas d'accident, la gravité sera certainement moindre pour le piéton (voir paragraphe 1.4) mais encore faut-il limiter la vitesse de l'impact voire si possible l'éviter.

Aujourd'hui, il semble que les progrès accomplis sur les systèmes de sécurité passive aient atteint un palier. En effet, après plusieurs années de développement, la sécurité passive est en train d'atteindre ses limites. Pour améliorer ostensiblement la protection des piétons, les constructeurs doivent travailler sur la sécurité active qui vise à éviter l'accident et non à l'accompagner. La sécurité active englobe l'ensemble des dispositifs permettant de prévenir les accidents (chocs, perte d'adhérence, décrochage...) par une action préventive (information du pilote-conducteur) et/ou dynamique (le dispositif « prend la main » dans le pilotage du véhicule). Schématiquement, ces dispositifs incluent un ensemble de capteurs (rotation des roues, glissement, accélération, proximité d'autres véhicules ou d'obstacles...), une chaîne de traitements des signaux émis par ces capteurs (circuit électronique, Asic, ordinateur de bord...), et un actionneur qui se substitue plus ou moins au pilote en prenant des actions correctives ou en générant une alerte (accélération, relâchement du freinage...). A titre d'illustration, les systèmes ABS (antiblocage des freins), ESP (correction de trajectoire) et antipatinage, équipent de plus en plus souvent les véhicules routiers en série. Afin d'améliorer la sécurité des vulnérables, les constructeurs cherchent actuellement à développer un logiciel permettant de détecter des obstacles à partir de capteurs comme les radars, lasers ou vidéos.

## 1.2.2 EuroNCAP

« **Comment réagirait votre voiture en cas d'accident ?** » Pour répondre à cette question, EuroNCAP <sup>5</sup>, une association indépendante, teste les différents véhicules mis sur le marché afin d'éclairer le consommateur dans son choix.

En effet, depuis 1996, EuroNcap pratique chaque année deux ou trois séries de *crash test* à la veille du printemps puis pendant l'été sur l'une des 6 grandes familles de véhicules (citadine, compacte, familiale, routière, petit monospace et grand monospace). En moyenne chaque année, elle teste une vingtaine de véhicules environ.

Alors que ces dernières années de plus en plus de véhicules obtenaient le score maximal de cinq étoiles, l'organisme indépendant EuroNCAP a décidé de rénover ses procédures de tests pour pousser les constructeurs à continuer à améliorer la sécurité de leurs véhicules. Toujours notés sur cinq étoiles, les résultats détaillés sont classés désormais selon quatre catégories : protection des adultes, des enfants, choc avec des piétons et

---

<sup>5</sup>European New Car Assessment Program

le nouveau "Safety Assist". Ce dernier vérifie maintenant les systèmes de sécurité embarqués tels que l'ESP (correcteur électronique de trajectoire appelé aussi ESC) bientôt obligatoire par ailleurs. Ces notes sont désormais données en pourcentage, afin de mieux apprécier les lacunes des différents véhicules testés, comme par exemple, la protection des piétons qui est souvent la plus faible.



FIG. 1.2 – Exemple de résultats EuroNcap du 18 février 2009 sur les derniers véhicules mis sur le marché - figure extraite de [www.euroncap.com](http://www.euroncap.com).

Désormais depuis 2009, les étoiles attribuées tiennent compte en plus de la protection habituelle des usagers du véhicule (adultes et enfants), de celle offerte aux piétons.

## 1.3 Les projets traitant la protection des piétons depuis 2002

La Commission Européenne est arrivée, le 11 juillet 2001, à un projet d'accord concernant un engagement volontaire de l'industrie automobile européenne destiné à accroître la protection des piétons impliqués dans des accidents avec des véhicules. La Commission a officiellement communiqué qu'il s'agissait d'une contribution déterminante et novatrice pour l'amélioration de la sécurité routière en Europe, ce qui est pour elle une priorité absolue. Cet engagement volontaire implique l'introduction, selon un planning ambitieux, d'une série de dispositifs de sécurité, tant sur le plan de la sécurité active que passive. Afin d'apporter des solutions concrètes à ce problème de sécurité, la Commission Européenne a cofinancé des projets européens tel que SAVE-U ou PReVENT. La France, de son côté, continue d'avoir une politique volontariste sur la sécurité routière, et finance à travers les

appels PREDIT <sup>6</sup>/ANR de nombreux projets tel que LOVE, CityVIP <sup>7</sup>...

### 1.3.1 SAVE-U : Capteurs et architecture système pour la protection des vulnérables



Lancé en mars 2002 [Meinecke *et al.*, 2005], son objectif principal est de développer un nouveau moyen de détection des piétons et des cyclistes fondé sur la fusion de trois technologies différentes de capteurs. En effet, il existe actuellement, des solutions pour détecter des piétons présents dans la trajectoire d'un véhicule, mais cette détection s'effectue au contact de la personne (sécurité passive).

L'avant du véhicule est alors prévu pour se déformer dès qu'il y a contact avec un objet afin de limiter les blessures pouvant être provoquées sur le piéton ou des « air bags » sont déployés à l'avant de la voiture afin d'amortir le choc sur la personne. Tous ces systèmes essaient de préserver le piéton sans éviter la collision.

L'objectif du projet SAVE-U est d'effectuer une détection à distance du piéton par analyse d'images visibles, thermiques et radar afin d'éviter la collision.

Le but du projet est de construire un prototype de faisabilité du système de détection de piétons. Le CEA-LIST <sup>8</sup>, partenaire du projet, mène ses recherches sur une architecture de traitement embarquée optimisée composée d'algorithmes de traitement d'images visibles et thermiques, d'une part, et, d'un modèle d'architecture reconfigurable, d'autre part.

Pour assurer une performance radicalement améliorée comparée à l'état de l'art, un concept innovant a été mis en oeuvre pour traiter les données image. En premier lieu, la segmentation d'images et les algorithmes de détection prennent en compte les informations provenant des deux caméras (IR et visible). Un microprocesseur de traitement d'images dédié a été conçu pour assurer le traitement en temps réel. En second lieu, de nouveaux algorithmes pour la classification d'usagers de la route non protégés, comme des piétons ou des cyclistes, ont été développés et mis en oeuvre. L'analyse est fondée sur des régions d'intérêt, pré-choisies, qui sont détectées par les algorithmes traitant l'image. Les algorithmes de classification prennent en compte de multiples facteurs, tels que la profondeur, le mouvement, la forme, la couleur ou la texture. La combinaison optimisée des deux traitements et la redondance résultant de l'utilisation des deux technologies différentes d'images assurent une détection et une classification des objets fiables.

<sup>6</sup>Programme national de REcherche et D'Innovation dans les Transports terrestres

<sup>7</sup>Déplacement sûr de véhicules individuels adaptés à l'environnement urbain

<sup>8</sup>Commissariat à l'Energie Atomique - Laboratoire d'Intégration des Systèmes et des Technologies

### 1.3.2 PReVENT : Applications de sécurité active et préventive



Ce programme [Fuerstenberg & Lages, 2005] supporté par la Commission Européenne vise à développer des systèmes à base de radars, de vidéo et de communication sans fil à l'horizon 2010 (date à laquelle Bruxelles souhaiterait pouvoir réduire de 50% la mortalité sur les routes par rapport aux chiffres de l'an 2000).

Lancé pour quatre ans en 2004, le programme PReVENT réunit les principaux constructeurs du secteur automobile (Audi, BMW, Fiat, PSA Peugeot Citroën, Renault, Volvo et Volkswagen), ainsi que des fournisseurs comme Bosch, Delphi, Philips, Siemens VDO, Navteq et Tele Atlas, sous la coordination de Daimler Chrysler. Onze projets ont été retenus, allant de la prévention des accidents par la détection des obstacles à l'assistance à la conduite en cas de perte de contrôle.

Des caméras sont par exemple installées en haut du pare-brise et de la vitre arrière ainsi que sur les rétroviseurs dans le projet APALACI de limitation des accidents.

Le projet SAFELANE permet d'alerter le conducteur s'il quitte sa trajectoire, mais aussi d'intervenir pour l'aider en agissant directement sur la direction. Des technologies qui devraient être disponibles dans un avenir « assez proche », selon des participants.

Leurs travaux ont fait l'objet de démonstrations pendant plusieurs jours en septembre 2007, sur les pistes de Satory, près de Versailles. En conditions réelles, les experts ont pu assister à des tests de freinage automatique, de détection d'obstacles et de communication de véhicule à véhicule. Le public a pu pour sa part approcher 25 véhicules expérimentaux.

### 1.3.3 Le projet LOVE

Le projet LOVE [Cornou *et al.*, 2008], est un projet ANR PREDIT avec un label SYSTEM@TIC<sup>9</sup> et MOVEo<sup>10</sup> pour la région francilienne. Comme il a été rappelé précédemment (cf. § 1.2.1), la législation évolue et les tests EuroNCAP intègrent aujourd'hui des tests de collision entre véhicules et piétons (sécurité passive). La prochaine étape est annoncée pour 2017 où les conditions de test (sécurité active) ne sont pas encore définies. Il est donc important pour les constructeurs et équipementiers français, en particulier pour Renault et Valeo (partenaires du projet), d'anticiper et de préparer ce qui pourrait être un système actif pour lequel une brique essentielle est la détection et la classification des vulnérables.

LOVe propose donc de contribuer à la sécurité routière en mettant principalement l'accent sur la sécurité des piétons. L'objectif est d'aboutir à des logiciels d'observation et de détection des piétons fiables et sûrs implantés sur des matériels compatibles avec une exploitation industrielle rapide. Pour ce faire, le projet propose de tester et de valider les solutions de dix laboratoires français (HeuDiasyC, IEF, CMM, CAOR, LIVIC, emotion, IMARA, ICARE, CEA-LIST et le LASMEA) qui sont spécialisés dans le domaine de la vision par ordinateur afin d'identifier quels sont les algorithmes qui permettent d'obtenir la meilleure solution possible de perception répondant aux exigences industrielles du projet.

---

<sup>9</sup>Pôle de compétitivité de la région Ile-de-France

<sup>10</sup>Pôle de compétitivité des automobiles et moyen de transports

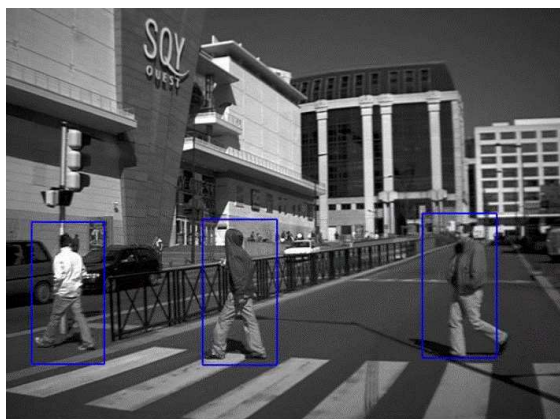


FIG. 1.3 – LOVE est un projet ANR PREDIT d'une durée de trois ans labellisé par Systém@tic et soutenu par MOV'EO et ViaMéca <sup>11</sup> - figures extraites de <http://love.univ-bpclermont.fr>.

### Renault



Ce projet doit permettre à Renault de se préparer aux nouvelles normes de sécurité active EuroNcap sur la protection des piétons ce qui lui confère un intérêt stratégique. En effet Renault pense que la mise au point de briques logicielles permettant de localiser les piétons doit lui permettre de conserver ses parts de marché en France et dans le monde en proposant des véhicules toujours à la pointe de la technologie.

Enfin ce projet doit lui permettre de disposer d'algorithmes qui pourront être adaptés à d'autres applications dans d'autres projets.

### Valeo



Valeo propose déjà aux constructeurs un module de sécurité active appelé Safe4U. La détection des piétons est effectuée par un radar placé sur la traverse supérieure et par deux caméras disposées le long de la calandre.

L'objectif principal pour lui est de réaliser à partir du projet LOVE, un nouveau système de détection de piétons pour les véhicules routiers. En effet, en tant qu'équipementier automobile, Valeo Vision souhaite apporter son savoir-faire en matière de spécification et de structure logicielle dans un environnement automobile.

Valeo permet au projet d'avoir des connaissances sur les contraintes industrielles de l'automobile tant sur la portabilité du système que sur le niveau de fiabilité à atteindre pour réduire effectivement et efficacement le nombre de morts grâce à ce système.

Ce projet doit permettre à Valeo d'obtenir des briques logicielles validées, standardisées et portables lui permettant par la suite de construire de nouveaux systèmes plus complexes afin de proposer un nouveau produit commercial.



### LASMEA



Le LASMEA (LABoratoire des Sciences des Matériaux pour l'Electronique et d'Automatique) est une entité mixte du CNRS et de l'Université Blaise Pascal de Clermont-Ferrand.

Dans le cadre du projet LOVE, le LASMEA va intervenir à quatre niveaux de la chaîne de perception/fusion en proposant :

- ▷ une définition de zones d'intérêts au sein desquelles la présence de piétons peut s'avérer dangereuse,
- ▷ une détection et un suivi des piétons à partir d'un capteur monovision,
- ▷ une détection et un suivi des piétons à partir d'un capteur stéréovision,
- ▷ une détection et un suivi des piétons à partir d'un capteur laser quatre plans,
- ▷ une fusion centralisée et décentralisée issus des différents systèmes de perceptions lasers et vidéos.

De cette collaboration le LASMEA se positionne sur un plan national dans le domaine des systèmes d'aide à la conduite et plus précisément sur la sécurité active en proposant des briques logicielles capables de détecter des piétons à partir de différents capteurs dans un contexte de conduite urbaine.

## 1.4 Contributions de la thèse

L'étude décrite dans cette thèse se situe dans le contexte du projet LOVE. Nous nous sommes donc particulièrement intéressés au développement d'un logiciel de détection et de suivi de piéton fiables à partir d'une séquences d'images lasers. Par la suite, nous nous sommes appliqués à traiter le problème de la fusion des données entre un capteur laser et un capteur vidéo.

L'étude d'un système de perception fondé sur un capteur laser a soulevé plusieurs problèmes théoriques récurrents. Dans un premier temps, le problème de détection : être capable de ne détecter que des piétons à partir d'une simple image de profondeur est loin d'être trivial surtout dans un environnement urbain. Dans un second temps, le problème de suivi et d'association de données : en effet, estimer et associer correctement des trajectoires incertaines comme les changements brusques de direction de piétons à partir de mesures provenant d'un ou plusieurs capteurs est un problème difficile à résoudre. L'efficacité d'une telle méthode dépend principalement de la (ou des) mesure(s) disponible(s) et des informations *a priori* connues (modèle dynamique, capteurs utilisés, bruits sur les mesures...). Partant de ce constat, les principales contributions de cette thèse s'organisent en deux parties :

- ▷ formulation d'une méthode de détection d'« objets-piétons » exploitant des données laser,
- ▷ formulation d'une méthode d'association de données pour filtre particulière.

### 1.4.1 Formulation d'une méthode de détection d'« objets-piétons » exploitant des données laser

La détection de piéton est un problème récurrent depuis de nombreuses années. Développer un algorithme efficace de détection puis de classification d'obstacles comme « objets-piétons » est compliqué quel que soit le contexte (intérieur ou extérieur) et le capteur utilisé (laser, video ou radar) par la grande variabilité du piéton en échelle, posture et apparence. Dans notre cas, détecter puis classifier un « objet-piéton » à partir d'un unique capteur laser peut même apparaître comme un challenge devant la relative pauvreté de l'information retournée par ce type de capteur. En effet, l'environnement observé par un capteur de profondeur se présente comme une collection de mesures de distance entre un référentiel connu (le véhicule) et les points des surfaces des objets placés dans la scène (cf. la figure 1.4). La liste de ces mesures (décrites sous forme de points) ne peuvent pas permettre avec certitude de classifier de manière simple et sûre chaque objet détecté par le télémètre. Ceci est d'autant plus vrai lorsque nous sommes dans un endroit aussi vaste et varié que l'environnement urbain où seuls les piétons doivent être détectés.

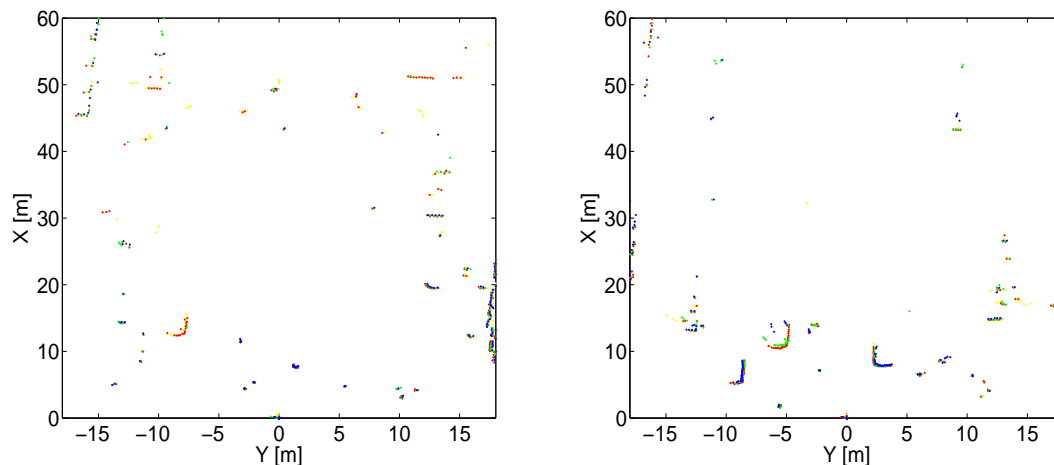


FIG. 1.4 – Exemple d'images laser enregistrée par un capteur laser quatre plans en contexte urbain. Les quatre couleurs permettent de distinguer les quatre plans du LIDAR.

La détection des piétons dans une image laser est traitée dans le chapitre 2 comme finalement un problème de classification d'obstacles en vision par ordinateur. En effet, un des problèmes habituels de la classification est la définition des contours de la classe qui décrivent au mieux les caractéristiques des objets recherchés dans l'image. Dans notre cas où l'on souhaite classifier deux classes (piéton, non piéton), les méthodes types machines d'apprentissage peuvent être utilisées, comme par exemple, les machines de type SVM <sup>12</sup> [Papageorgiou & Poggio, 2000]. Habituellement les machines d'apprentissage SVM utilisent une base d'exemples positifs et d'exemples négatifs. Malheureusement dans notre cas, la construction de la base d'exemple est délicate voire impossible pour deux raisons principales : (1) la grande variabilité de forme de l'objet et (2) la pauvreté de l'information délivrée. Pour répondre à ce problème, nous proposons d'extraire les « objets-piétons » à partir d'une méthode non paramétrique fondée sur les méthodes à

<sup>12</sup>Support Vector Machine

noyau (KDE <sup>13</sup>) qui permet de modéliser la classe piéton par une densité de probabilité. En effet, cette méthode a l'avantage de construire à chaque nouvelle image laser et sans connaissance *a priori* du nombre d'obstacles présent dans la scène observée, la fonction de vraisemblance de la classe piéton à partir de ses caractéristiques géométriques propres.

Dans un second temps, cette thèse propose d'étendre notre approche à l'analyse des quatre plans du capteur laser. Nous proposons un algorithme capable de prendre une décision finale sur la classification ou non d'un objet comme piéton, qu'après avoir analysé tous les objets classifiés par l'étape précédente dans les quatre nappes. L'idée basique est d'améliorer les performances des algorithmes de détection de piéton fondés sur un seul plan laser, et de compenser les mauvaises détections qui peuvent apparaître de manière ponctuelle dans un plan par les autres plans. En effet, on peut légitimement penser que cette méthode va permettre de diminuer le nombre de fausses alarmes qui apparaissent de manière aléatoire dans une nappe mais jamais dans les quatre plans à la fois. De plus, on peut espérer améliorer le taux de détection de piéton en vérifiant la présence d'« objets-piétons » dans les quatre nappes laser au lieu d'une seule. Cet algorithme de fusion repose également sur une méthode non paramétrique fondée sur les méthodes à noyau permettant de modéliser le nombre de nappes d'un capteur laser multi-plans.

### 1.4.2 Formulation d'une méthode d'association de données pour filtre particulaire

Dans un second temps, cette thèse s'intéresse à la fusion (ou combinaison) de données entre un capteur laser et un capteur vidéo. L'objectif est d'obtenir un pistage robuste des piétons situé dans un environnement routier complexe tel que l'environnement urbain. Une étude sur la détection des piétons à bord de véhicule routier présentée dans [Gavrila *et al.*, 2001], indique que la coopération entre un capteur laser et une caméra apparaît comme une des meilleures solution actuelles. La fusion des informations délivrées par les deux systèmes doit permettre d'utiliser opportunément les différentes caractéristiques complémentaires et redondantes de manière à améliorer le taux de détection et de diminuer le nombre de fausses alarmes. Le problème spécifique du suivi de trajectoires complexes multi-pistes multi-capteurs telles que l'interaction de plusieurs objets dans l'espace capteur (cf. la figure 1.5) est traité durant le chapitre 3.



FIG. 1.5 – Exemple d'un suivi de piéton avec deux caméras. Gauche : caméra 1, milieu : caméra 2, droite : trajectoires en vue de dessus - figure extraite de [Bardet, 2009]

<sup>13</sup>Kernel Density Estimation



Afin de résoudre ce problème, nous avons essayé de répondre aux questions suivantes :

- ▷ Quel type de modèle probabiliste doit-on utiliser et/ou développer à partir de l'état de l'art afin d'améliorer le suivi multi-piétons ?
- ▷ Comment associer au mieux les nombreuses détections provenant des différents capteurs et des différentes cibles observées ?

Ainsi, dans le cadre spécifique du suivi multi-objets à partir d'un capteur laser et d'un capteur vidéo, cette thèse propose un algorithme non paramétrique d'association de données fondé sur le filtrage particulaire. Cette méthode permet de mieux prendre en compte toute l'information disponible entre les pistes (informations connues *a priori*) et les observations (provenance, confiances et précisions) .

## 1.5 Organisation du document

Ce document est organisé en quatre parties selon le plan suivant (cf. la figure 1.6) :

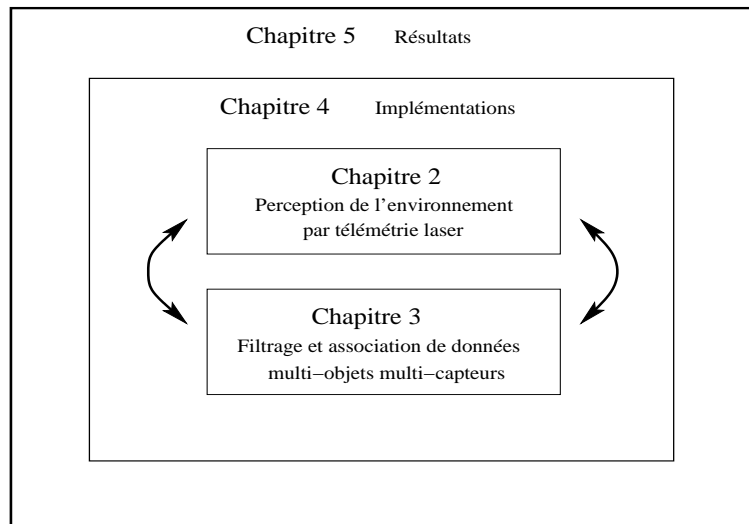


FIG. 1.6 – Vue d'ensemble de la thèse.

### Chapitre 2 : Modélisation de la perception de l'environnement par télémétrie laser

Ce chapitre présente un rapide état de l'art technologique des capteurs laser disponibles sur le marché afin d'identifier le télémètre laser qui apparaît le mieux adapté à nos travaux. Différentes techniques de segmentation ainsi que des méthodes de détection/reconnaissance d'obstacles à partir d'une image de profondeur sont présentées. Ensuite, et afin de positionner les travaux de cette thèse, un état de l'art sur les différentes méthodes existantes permettant de détecter des piétons à partir d'un capteur laser est proposé. Pour conclure ce chapitre, les particularités géométriques qui caractérisent un piéton dans une image laser sont listées et détaillées.

### Chapitre 3 : Filtrage et association de données multi-objets multi-capteurs

Ce chapitre commence par présenter le principe général du filtrage bayésien. Nous rappelons ensuite les différents algorithmes de filtrage bayésien, à savoir le filtre de Kalman

et ses extensions, ainsi que les méthodes de Monte Carlo tels que les filtres particulaires et MCMC (Monte Carlo Markov Chain). Chaque méthode est commentée en montrant les avantages mais aussi leurs limites. Ce chapitre se termine sur l'extension du problème de suivi multi-objets au cas multi-capteurs en proposant dans un premier temps, un état de l'art des principales méthodes existantes avant de terminer sur les différents types d'architectures qui peuvent être rencontrés.

### **Chapitre 4 : Implémentations**

Le chapitre 4 est dédié aux algorithmes de détection et de suivi des piétons mono et multi-capteurs qui ont été développés durant cette thèse. Dans un premier temps, l'algorithme de détection non paramétrique de piétons basé sur l'utilisation d'un capteur laser quatre nappes est détaillé. Dans un second temps, l'algorithme particulière fondé sur une méthode d'association de données non paramétrique qui permet une association non-déterministe des données est présenté. Cet algorithme très général a été étendu au cas multi-capteurs.

### **Chapitre 5 : Résultats**

Enfin le chapitre 5 présente de nombreux résultats simulés et expérimentaux effectués sur des données réelles enregistrées à bord du véhicule VIPA du LASMEA et du véhicule expérimental de Renault. Des critères sont proposés pour quantifier les performances de détection et de suivi des piétons. Les résultats montrent que l'utilisation des quatre plans améliore le taux de détection et diminue significativement le nombre de fausses alarmes. Les meilleurs résultats sont obtenus lorsque nous fusionnons des détections laser avec des détections vidéo.

Enfin, la conclusion générale donne une synthèse de cette étude et quelques perspectives de développement de ses travaux.



---

## Chapitre 2

# Perception par télémétrie laser

Dans ce chapitre, après avoir rappelé le principe de la télémétrie laser, nous nous sommes intéressés aux spécifications des principaux capteurs existants en nous interrogeant sur le choix qui permet une industrialisation automobile viable. Ce choix qui détermine la densité d'acquisition de l'environnement influera par la suite sur les méthodes et les performances que nous pourrions espérer de l'algorithme de détection. Par la suite un rappel des principales techniques de segmentation ainsi que des méthodes de détection/reconnaissance d'obstacles à partir d'une image de profondeur est présenté. Puis, nous dressons un panorama des différentes méthodes proposées ces dernières années pour détecter des piétons à partir d'un capteur laser. Ensuite, nous proposons une définition d'un détecteur de piétons « idéal » avant d'expliquer pourquoi la classification d'« objets-piétons » est difficile en perception par ordinateur et particulièrement avec un capteur laser. Ce chapitre se termine en précisant quels sont les piétons à détecter dans le cadre de nos travaux c'est-à-dire en fonction du cahier des charges fixé par le projet LOVE.

## 2.1 Principe de la télémétrie laser à balayage

### 2.1.1 Généralité

Un télémètre laser permet de réaliser une mesure précise entre l'instrument (le capteur) et un objectif. Ce système de mesure, également appelé LIDAR (Light Detection And Ranging) ou IRDAR (Infra-Red Detection And Ranging) est conçu à l'aide de lasers émettant respectivement dans le visible et dans l'infra-rouge. Les qualités de cohérence spatiale et temporelle du rayonnement permettent des divergences très faibles du faisceau et des focalisations poussées. Une autre caractéristique importante qui est liée à l'utilisation de tels capteurs, est la réponse optique des cibles. La réémission ou la diffusion du rayonnement incident est un phénomène qui dépend de nombreux paramètres de nature et d'état des corps en question (notamment les propriétés de structure). Selon la rugosité de la surface éclairée, on observe une réflexion diffuse pour une surface rugueuse ou spéculaire pour une surface lisse. L'énergie reçue peut être réfléchie, transmise d'un milieu à un autre, ou absorbée par des surfaces sombres par exemple. Enfin, des normes sévères sont à respecter en vue d'une utilisation spécifique. La portée des lasers est limitée, dans les applications civiles, par des considérations de sécurité oculaire.

Par suite de ses propriétés de cohérence et des puissances qu'il est capable d'émettre

dans des intervalles de temps réglables, le laser a suscité des développements importants dans le domaine de la télémétrie. Selon le procédé technique mis en oeuvre et l'ordre de grandeur des distances à mesurer, il est possible de distinguer quatre méthodes assez différentes. Ainsi, la mesure de distance peut être obtenue par télémétrie impulsionnelle, mesure par comparaison de phase, mesure par modulation de fréquence, voire triangulation active. Les trois premières méthodes permettent d'accéder directement à l'information distance alors qu'elle n'est obtenue que de manière indirecte dans le cas de la triangulation active. Compte tenu du contexte applicatif qui nous motive dans cette thèse, nous rappelons la première technique également nommée « mesure par temps de vol ». C'est sur ce principe que fonctionnent les capteurs décrits [Fuerstenberg *et al.*, 2002, Fuerstenberg *et al.*, 2003, Fuerstenberg & Dietmayer, 2004, Fuerstenberg & Lages, 2005] et utilisés dans cette thèse. Pour une information complète sur les procédés d'acquisition de données tridimensionnelles, le lecteur pourra consulter l'ouvrage [Gallice, 2002].

### 2.1.2 Du capteur à l'image

Un télémètre laser de type LIDAR est un appareil permettant de mesurer les distances entre l'environnement (à une distance comprise entre quelques dizaines de centimètres et une centaine de mètres) et le capteur dans une direction donnée. C'est à partir d'une impulsion, de très courte durée mais de grande puissance, émise par le LIDAR et qui est dirigée vers un objectif qui permet de déterminer sa distance avec le capteur. En effet, l'écho provenant de la diffusion ou de la réflexion de l'onde électromagnétique sur la cible est capté par un détecteur photoélectrique. La distance entre le télémètre et l'objectif est alors déduite à partir de la mesure précise de l'intervalle de temps séparant l'émission de l'impulsion laser et de la réception du signal de retour où il suffit de connaître la vitesse de propagation du rayonnement. La relation qui lie le temps à la distance est simple :  $v \cdot \tau = 2 \cdot \rho$  où  $v$  est la vitesse de propagation du signal,  $\rho$  la distance à l'objet réfléchissant et  $\tau$  le temps de transit du signal séparant l'émission de la réception, appelé temps de vol.

Pour l'instant, la technique de mesure décrite ne nous permet d'obtenir qu'une information ponctuelle de distance. Pour former une ligne (ou nappe, ou profil de distance) voire une image, il est nécessaire de dévier le faisceau émis par le capteur en azimut (angle que fait le plan vertical passant par l'axe de visée et le plan vertical de référence) voire en site (angle que forme, avec sa projection dans le plan horizontal, une ligne joignant un observateur à un point visé). Les solutions les plus fréquemment employées font intervenir un prisme tournant ou un jeu de miroirs.

Pour surmonter des problèmes de temps d'acquisition trop lent, un imageur de distance doit être doté d'un système de balayage performant. Un tel système implique des montages mécaniques de déflexion complexes et par conséquent très onéreux. Remarquons que la précision angulaire avec laquelle le faisceau est dirigé se répercute directement sur la précision de mesure. Ainsi à titre d'exemple, nous pouvons indiquer que les capteurs décrits dans [Fuerstenberg *et al.*, 2002, Fuerstenberg *et al.*, 2003, Fuerstenberg & Dietmayer, 2004, Fuerstenberg & Lages, 2005] fournissent des informations avec une précision en distance de l'ordre du centimètre et une résolution angulaire azimutale variant de  $0.25^\circ$  à  $1^\circ$  selon la fréquence de rotation du prisme. L'échantillonnage obtenu est de type incrément angulaire constant.

Enfin, la représentation des données de profondeur dépend directement du système de déflexion du télémètre laser et donc de sa géométrie. Les données issues d'un cap-

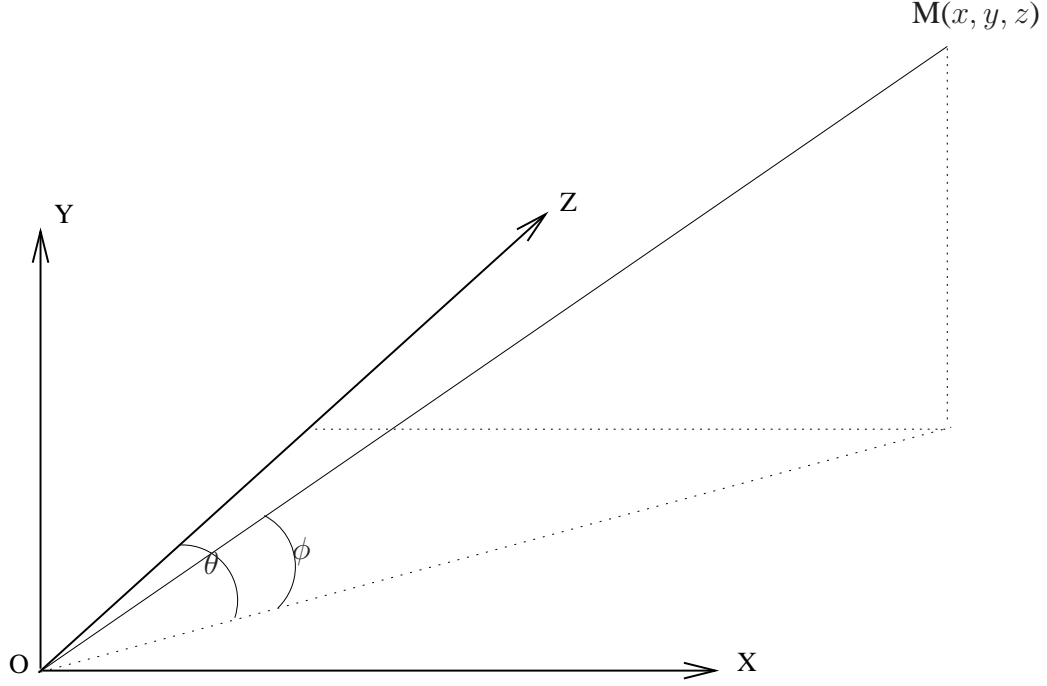


FIG. 2.1 – Système de coordonnées cartésiennes et pseudo-sphériques.

teur de profondeur se présentent comme une collection de mesures de distance entre un référentiel connu et les points des surfaces des objets placés dans la scène. La liste de ces mesures peut être écrite sous la forme  $xyz$  ;  $x$ ,  $y$  et  $z$  sont les trois coordonnées d'un point exprimées dans un système d'axes orthogonaux (cf. la figure 2.1). Bien que d'autres géométries de capteurs optiques puissent être trouvées, la plupart des capteurs de profondeur exploités dans la synthèse bibliographique qui suit, comme ceux que nous utilisons (cf. § 2.1.4), se modélisent par un système de coordonnées sphérique ou plus exactement « pseudo-sphérique »<sup>1</sup>.

En posant que la mesure repérée par son indice ligne  $i$  et colonne  $j$  est  $\rho_{ij}$ , les coordonnées  $(x, y, z)$  d'un point  $M$  de l'espace sont données par les équations générales :

$$\begin{cases} x = a_x + s_x \rho_{ij} \cos(i\Delta\phi) \sin(j\Delta\theta) \\ y = a_y + s_y \rho_{ij} \sin(i\Delta\phi) \\ z = a_z + s_z \rho_{ij} \cos(i\Delta\phi) \cos(j\Delta\theta) \end{cases} \quad (2.1)$$

où  $\mathbf{a} = (a_x, a_y, a_z)^T$  représente le possible vecteur translation entre l'origine du capteur et l'origine du repère, et où  $\mathbf{s} = (s_x, s_y, s_z)^T$  représente les pas d'échantillonnage ou encore les facteurs d'échelle.  $\Delta\theta$  et  $\Delta\phi$  correspondent respectivement aux incréments angulaires dans la direction horizontale et verticale le cas échéant.

### 2.1.3 Bilan des capteurs LIDAR exploitables

Dans un contexte automobile, le coût des capteurs est un souci majeur des constructeurs, car même si le consommateur veut toujours plus de sécurité, il veut également des véhicules toujours moins chères. A partir de ce constat, les constructeurs se doivent d'améliorer leurs véhicules sans en augmenter le prix. Choisir le meilleur capteur au

<sup>1</sup>On parle plus généralement dans un système de coordonnées sphérique de longitude et de co-latitude.

meilleur prix est donc une tâche primordiale. En effet, ce choix influera directement sur les méthodes envisagées et donc sur les algorithmes développés.

Historiquement, le LASMEA travaille depuis longtemps sur le développement de logiciels de perception à partir de capteurs lasers tels que le Riegl LD90-25S [Checchin, 1996], le Riegl LMSZ210-60 [Blanc, 2005], le Sick LMS 221 [Chanier *et al.*, 2008] et depuis 2006 sur l'IBEO LD ML [Gidel *et al.*, 2008a] qui équipe actuellement le véhicule VELAC du laboratoire (cf. le tableau 2.1). On trouvera en annexe 1, un tableau présentant les principaux capteurs laser 2D existant sur le marché.

Référence capteur	Propriété
Sick LMS 221	Mono-nappe
IBEO LDML/alasca XT	4 nappes
Riegl LMSZ210-60	1666 nappes

TAB. 2.1 – Etat de l'art des capteurs disponible au LASMEA et/ou dans le contexte du projet LOVE.

### Solutions exploitant une nappe laser



Sick LMS 221

Avec ce type de capteur, on distingue deux types d'implantation. La première exploite une nappe laser 2D montée horizontalement sur le véhicule. Avec un capteur tel que le Sick LMS 221 on obtient un champ de vue maximal de  $180^\circ$  (H) soit 720 impacts laser maximum avec une résolution de  $0.25^\circ$  (cf. la figure 2.2). Elle est principalement utilisée dans des applications de détection d'obstacle [Diosi & Kleeman, 2003] ou pour reconstituer la dimension manquante (profondeur) dans des applications de fusion laser/vidéo [Gidel *et al.*, 2009b].

Le principal inconvénient de cette méthode réside dans le fait que l'information de profondeur est accessible uniquement dans un plan de l'image vidéo. Cela implique l'impossibilité d'obtenir une information sur la hauteur des objets détectés ainsi qu'une certaine sensibilité à l'occultation d'objets entre-eux. La seconde implémentation consiste en une nappe laser 2D montée autour d'un axe de rotation vertical permettant d'obtenir un champ de vue maximal de N degrés en horizontal (N étant fonction de la rotation maximale en azimut du montage) et de  $180^\circ$  en vertical soit  $720 \cdot N$  impacts laser avec une résolution de  $0.25^\circ$ . Ce montage est adopté par exemple par l'université de Hanovre [Wulf & Wagner, 2003] (cf. la figure 2.3). Même si l'idée est intéressante, elle reste inapplicable dans un contexte routier. En effet, le système a besoin de  $15ms$  pour obtenir un plan laser. En faisant l'hypothèse que nous souhaitons un angle d'ouverture horizontal de  $N = 60^\circ$  avec une résolution de  $0.25^\circ$ , le système aura besoin de  $0.015 \cdot 60 \cdot 4 = 3.6$  secondes.



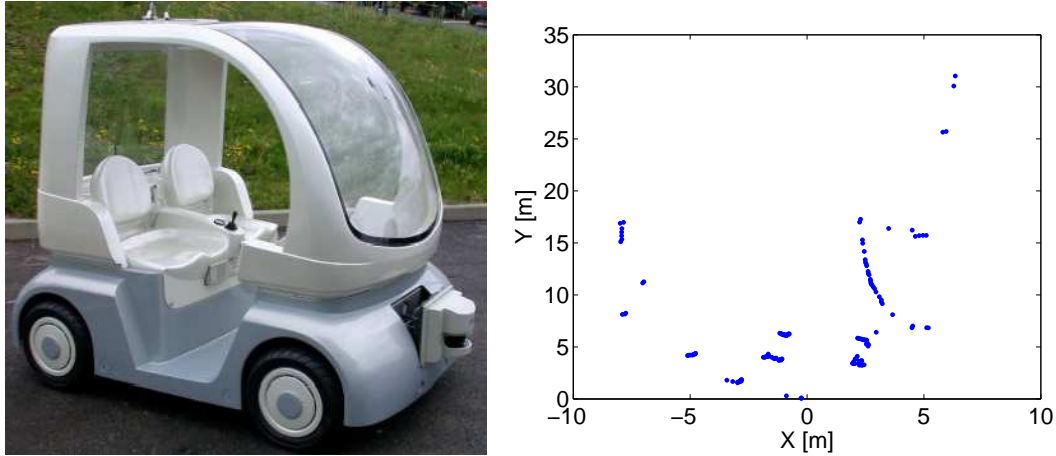


FIG. 2.2 – A gauche : un véhicule VIPA du LASMEA équipé d'un capteur Sick. A droite : un plan laser obtenu avec le capteur Sick.

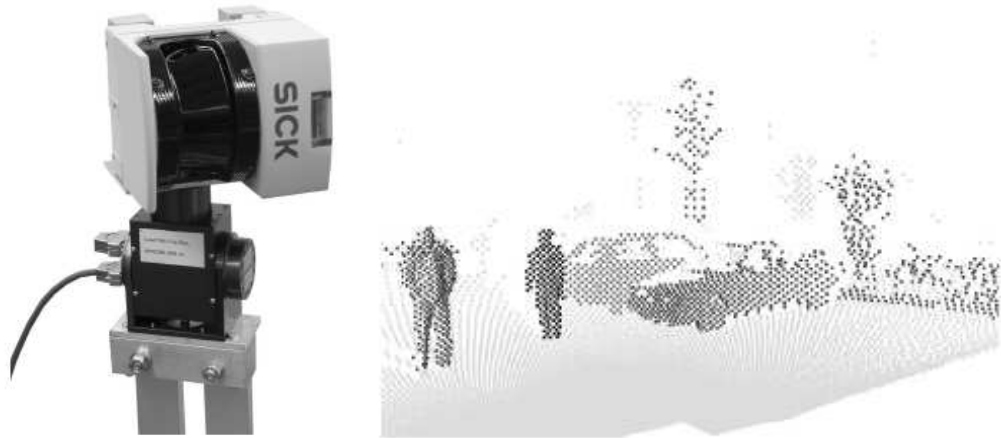


FIG. 2.3 – A gauche : le capteur Sick monté autour d'un axe de rotation vertical. A droite : un résultat obtenu à partir d'un capteur Sick en rotation azimuth - figure extraite de [Wulf & Wagner, 2003].

### Solutions exploitant plusieurs nappes laser



IBEO LD ML

Avec ce type de capteur, on obtient une juxtaposition verticale de plusieurs nappes en faisant appel à des produits tel que les modèles multi-plans à 4 faisceaux de chez IBEO. L'IBEO LD ML ou l'IBEO alasca XT proposent de scanner simultanément l'environnement sur 4 plans laser horizontaux séparés verticalement d'environ  $1.1^\circ$  chacun. Ainsi, avec une résolution de  $0.25^\circ$  et un angle d'ouverture de  $180^\circ$ , le capteur peut fournir 2880 impacts laser soit 720 impacts par plan.

Le lecteur peut consulter les publications suivantes [Fuerstenberg *et al.*, 2002, Fuerstenberg *et al.*, 2003, Fuerstenberg & Dietmayer, 2004, Fuerstenberg & Lages, 2005] pour tous les détails concernant ces deux capteurs. Ces produits sont déjà très utilisés par les



laboratoires (INRIA <sup>2</sup>, HeuDiaSyC <sup>3</sup>, LASMEA) car ils sont dédiés à une utilisation pour l'automobile. Le véhicule VELAC du LASMEA est actuellement équipé d'un capteur IBEO LD ML (cf. la figure 2.4)

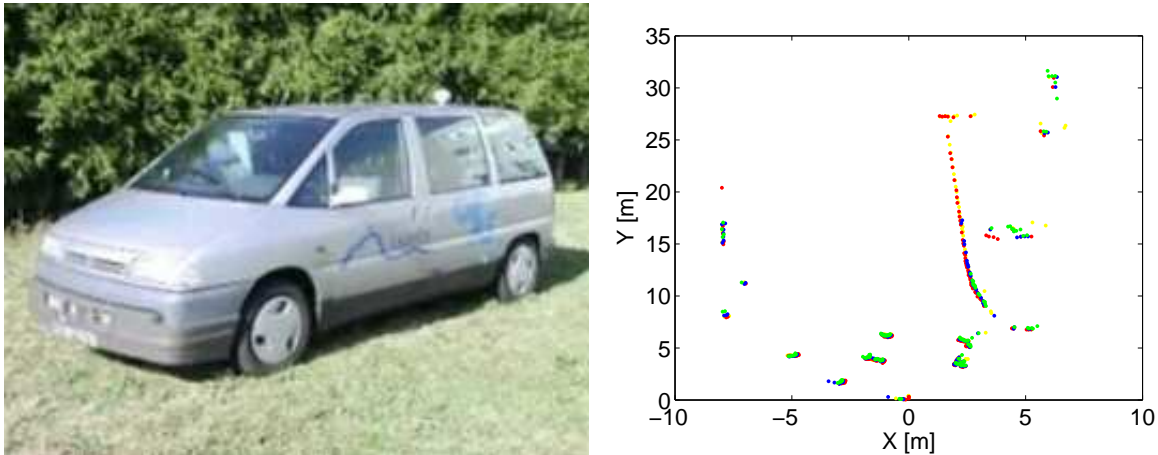


FIG. 2.4 – A gauche : le véhicule VELAC. A droite : un résultat obtenu à partir d'un capteur IBEO 4 plans. Chaque couleur correspondant à un plan laser.

### Solutions exploitant un imageurs laser 3D



Riegl LMSZ210-60

Les télémètres laser 3D (ou imageurs 3D) constituent certainement le capteur le plus attendu par les roboticiens abordant la robotique mobile en milieu extérieur. Par exemple le Riegl LMSZ210-60 permet d'acquérir une véritable information de profondeur sur un secteur étendu et non plus sur une coupe comme dans le cas d'un capteur mono-nappe (cf. la figure 2.5). Il permet notamment d'obtenir une image sur un champ de vue de  $60^\circ$  en vertical avec une résolution de  $0.036^\circ$  par  $333^\circ$  en horizontal avec un angle de résolution de  $0.018^\circ$ .

Les applications ciblées pour un tel instrument sont très variées telle que la visualisation détaillée de structures de génie civil, la reconstruction de site, monument ou artefact [Fournier *et al.*, 2006]. Les principaux défauts de ce type de capteur sont son prix élevé et sa cadence d'acquisition trop lente (plusieurs secondes pour obtenir une acquisition comme dans la figure 2.5).

#### 2.1.4 Capteur retenu

Au vu de l'état de l'art technologique donné dans le paragraphe précédent, les partenaires industriels ont choisi d'utiliser, dans le cadre du projet LOVE, le télémètre laser quatre nappes IBEO ALASCA XT (cf. la figure 2.6).

<sup>2</sup>Institut National de Recherche en Informatique et Automatique

<sup>3</sup>Heuristique et Diagnostic des systèmes Complexes de l'université de technologie de Compiègne



FIG. 2.5 – Image d'intensité dense d'un bureau - figure extraite de [Blanc, 2005].

Ce capteur développé par la société IBEO entre dans la catégorie des lasers multi-plans et propose de scanner l'environnement sur quatre plans répartis sur un angle d'ouverture approximatif de  $3.2^\circ$  (cf. la figure 2.7) et fourni des performances assez similaires à l'IBEO LDML installé sur le véhicule du LASMEA. En effet, ils possèdent tous les deux un champ de vue pouvant atteindre  $270^\circ$ . Le constructeur indique une portée maximale de 128 m pour l'IBEO LD ML et de 256 m pour l'ALASCA XT avec une précision de  $\pm 5$  cm. La résolution angulaire est, quant à elle, fonction de la fréquence d'acquisition, qui peut varier de 0.25° à 10Hz jusqu'à  $1^\circ$  à 40Hz.



FIG. 2.6 – A gauche, le capteur IBEO LD ML qui équipe le véhicule VELAC du LASMEA. A droite, le capteur IBEO ALASCA XT qui équipe un des véhicules tests de Renault.

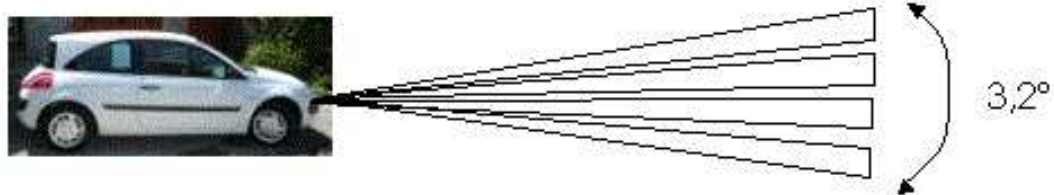


FIG. 2.7 – Capteur laser multi-plans avec quatre plans répartis sur un angle d'ouverture de  $3.2^\circ$ .

Dans cette thèse, le capteur a été monté de telle sorte que sa deuxième nappe soit parallèle au sol [Gidel *et al.*, 2009a]. En effet, si le capteur était positionné tel que la première nappe soit parallèle au sol, la quatrième nappe ne pourrait alors détecter que de "grands" piétons au delà de 20 m et plus rien au delà de 30 m (cf. la figure 2.8). De plus dans cette configuration, la première nappe peut être utilisée afin de surveiller le

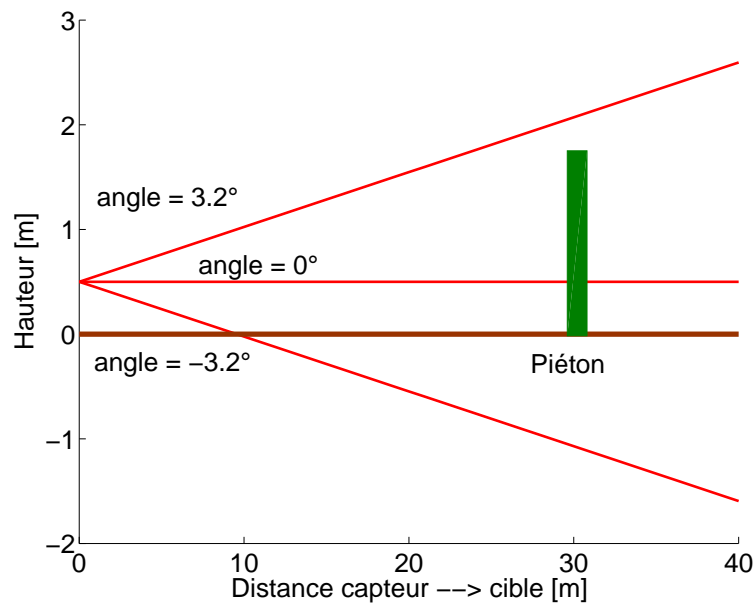


FIG. 2.8 – Position d’un plan laser en fonction de l’angle de tangage du véhicule. Comme le montre cette figure, un piéton localisé à 30 m ne peut pas être détecté avec un angle de tangage de  $3.2^\circ$ .

tangage du véhicule. Le capteur IBEO alsca XT a notamment permis de réaliser la campagne d’enregistrement des scénarios piétons en ville à bord d’un véhicule expérimental appartenant à Renault. Le capteur IBEO LD ML a quant à lui réalisé une campagne d’enregistrement de plusieurs scénarios piétons sur le campus d’Aubière autour du laboratoire du LASMEA.

### 2.1.5 Conclusion

Leur excellente précision en mesure de distance (avec un bruit de l’ordre de quelques centimètres) à une distance relativement grande (plusieurs dizaines de mètres) et leur bonne résolution angulaire (avec une mesure de distance tout les dixièmes de degrés) sur une zone pouvant atteindre 270 degrés font des télémètres laser à balayage des dispositifs adaptés pour un relevé géométrique précis et dense de l’environnement. Toutefois, dans un contexte industriel où le prix des capteurs est très important, cette dernière remarque est à tempérer lorsque les capteurs utilisés sont « bon marché » et propose une résolution angulaire de seulement un degré. En effet, à une distance de 30 mètres, ce type de capteur ne recueille plus qu’un impact tout les 52 cm, limitant de fait, la taille des objets détectés.

De plus, il faut garder à l’esprit que ce type de capteur optique est sensible aux conditions météorologiques, par nature changeantes en extérieur. La pluie, la neige et le brouillard affectent la portée et la précision de la mesure de distance. Par ailleurs, l’électronique embarquée de traitement du signal nécessaire sur ces capteurs ainsi que le dispositif mécatronique de balayage rendent leur coût très élevé.

## 2.2 Méthodes de détection/reconnaissance d'obstacles

Dans un relevé télémétrique, tous les points qui constituent l'image de profondeur appartiennent à un obstacle. Le processus de détection/reconnaissance d'obstacles consiste alors à détecter les discontinuités entre 2 points successifs afin que les obstacles apparaissent comme des ensembles de points voisins avec une même étiquette. Cette difficile tâche s'apparente donc à un processus de segmentation de l'image laser (cf. la figure 2.9). En effet, il s'agit bien de partitionner l'image initiale en régions homogènes sur la base de certains critères afin d'extraire une information plus riche, permettant par la suite une interprétation de l'image plus simple à manipuler.

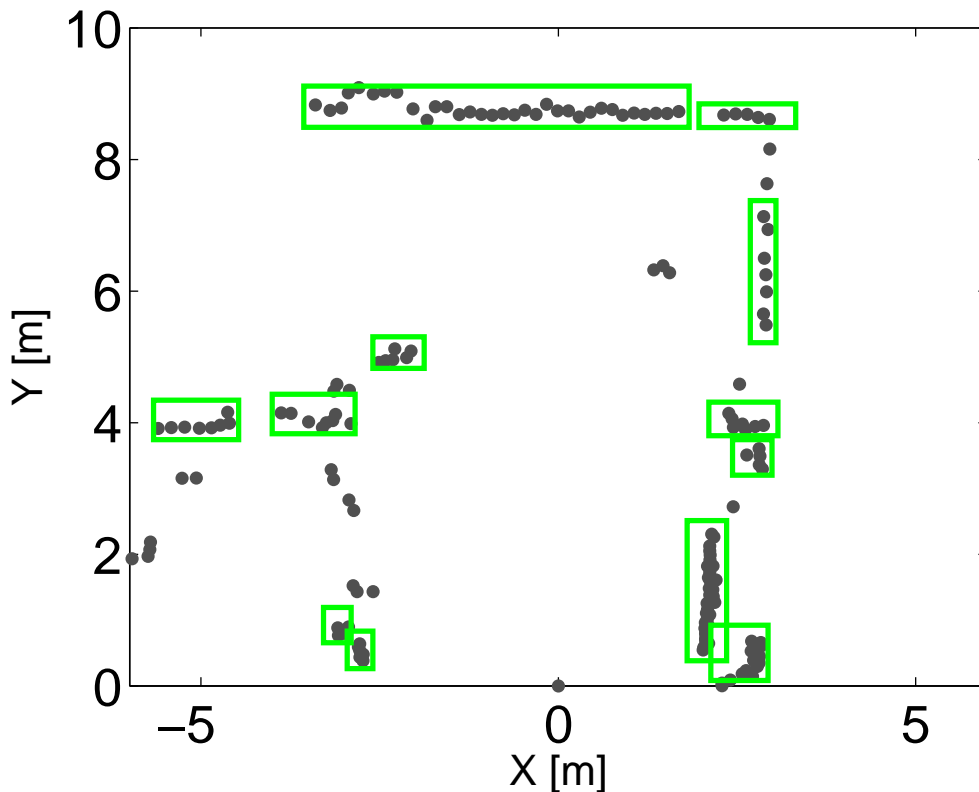


FIG. 2.9 – Exemple de segmentation d'une image d'impacts laser en vue de dessus. Un groupe de points doit présenter au minimum trois points consécutifs.

Ce problème de segmentation peut être résolu comme un problème de suivi [Bardet, 2009] à travers deux stratégies illustrées par la figure 2.10 :

- ▷ l'approche directe,
- ▷ l'approche inverse.

### 2.2.1 Méthodes directes

Ces stratégies consistent à construire une observation à partir d'un état (carte de l'environnement ou gabarits d'objets *a priori* connus). On ne cherche donc pas à expliciter la fonction de transformation inverse  $\mathbf{X} = g(\mathbf{Z})$ , mais on cherche la transformation  $\mathbf{Z} = f(\mathbf{X})$  qui permet de faire correspondre une observation avec une prédiction de l'état

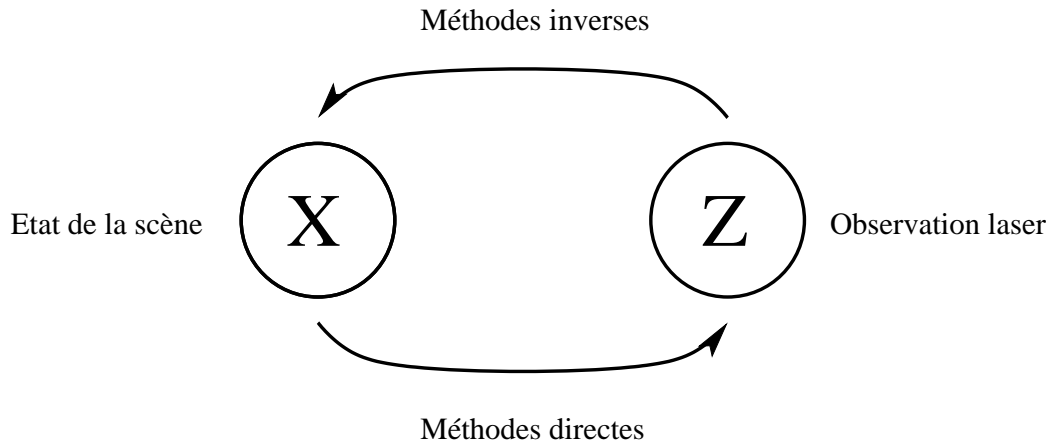


FIG. 2.10 – Schéma des deux approches permettant d’estimer l’état  $X$  qui représente notre compréhension de la scène. L’estimation de l’état à partir d’une observation capteur est un problème inverse. Le processus qui consiste à construire une observation à partir d’un *a priori* sur la carte ou sur les objets cherchés est un problème direct.

$X$ . On est alors comme pour un problème de suivi, dans une approche basée modèle qui demande une connaissance sur : le modèle de l’objet, le modèle de l’observation qu’il génère et une fonction de vraisemblance mesurant la concordance entre l’observation  $Z$  et l’état prédit  $X$ .

Ces méthodes sont principalement utilisées dans des application de navigation comme dans [Cox, 1991] et [Gutmann & Konolige, 1999] où les auteurs utilisent un algorithme qui fait la correspondance entre les points laser (cf. la figure 2.11) et une carte *a priori* connue par la méthode de minimisation des moindres carrés. La perception de l’environnement peut également être faite à partir de méthodes venant de la vision par ordinateur comme l’algorithme ICP (Iterative Closest Point) [Besl & McKay, 1992] qui permet de faire la correspondance entre les points de deux images laser successives. Dans des applications de localisation dans un environnement inconnu, d’autres auteurs, comme [Lu & Milios, 1997] ou [Gutmann & Konolige, 1994], proposent de faire correspondre entre deux images laser consécutives les points 2D, qui minimisent une certaine distance obtenue par odométrie.

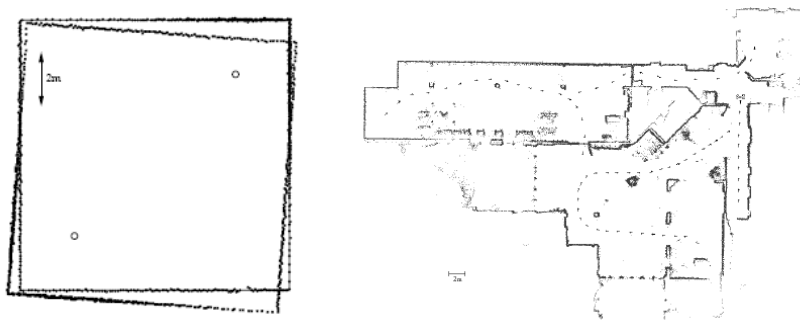


FIG. 2.11 – Principe d’une mise en correspondance entre deux images laser consécutives. A gauche, deux images laser consécutives sans mise en correspondance A droite, résultat après une mise en correspondance - figure extraite de [Gutmann & Schlegel, 1996].

### 2.2.2 Méthodes inverses

Ces stratégies consistent à construire l'état  $\mathbf{X}$  d'une scène réelle à partir d'observation  $\mathbf{Z}$ , en appliquant la transformation  $\mathbf{X} = g(\mathbf{Z})$ . Elles fournissent alors directement une estimation de l'état à partir des observations, sans connaissance aucune de la scène observée. Ces approches sont couramment utilisées pour la détection d'obstacle, où l'état  $\mathbf{X}$  se réduit en une variable binaire définissant la présence ou non d'objet.

La plupart de ces méthodes regroupent préalablement les différents points 2D en cluster, avant d'essayer de les classer sous plusieurs formes géométriques simples comme les segments, lignes, cercles ou amas de points...

On rappelle que l'objectif du clustering est de regrouper dans chaque image laser, les points consécutifs afin de répertorier l'ensemble des objets présents. Pour cela, il faut calculer la distance qui sépare un point  $P_i$  de son voisin immédiat  $P_{i+1}$ . Parmi les nombreuses formes géométriques, la classification des clusters en lignes ou segments est la plus répandue (cf. la figure 2.12). On peut citer l'algorithme "Split-and-Merge", qui est probablement l'algorithme d'extraction de ligne le plus populaire [Borges & Aldon, 2004, Siadat *et al.*, 1997, Zhang & Ghosh, 2000]. Dans d'autres applications [Taylor & Probert, 1996, Vandorpe *et al.*, 1996], l'algorithme « Incremental » est utilisé pour sa simplicité alors que l'algorithme "RANSAC" [Fisler & Bolles, 1981] présente l'avantage d'être assez générique pour être utilisé dans de nombreuses situations.

D'autres travaux comme ceux mentionnés dans [Premebida & Nunes, 2005] proposent de classer ces clusters en cercle ou en ellipse en proposant de résoudre l'équation du cercle ou de l'ellipse par une décomposition SVD. Dans [Xavier *et al.*, 2004], la classification des clusters en cercle utilise une nouvelle technique appelée Inscribed Angle Variance (IAV) (cf. la figure 2.12). Enfin [Arras & Vestli, 1998] proposent de classer les clusters comme des amas de points qui contiennent un certain nombre de points 2D dans un espace donné.

**Notre algorithme de segmentation repose sur une stratégie inverse, fondé sur l'algorithme développé par [Arras, 2003] qui propose de construire les différents clusters sous deux formes géométriques : des segments et des amas de points.**

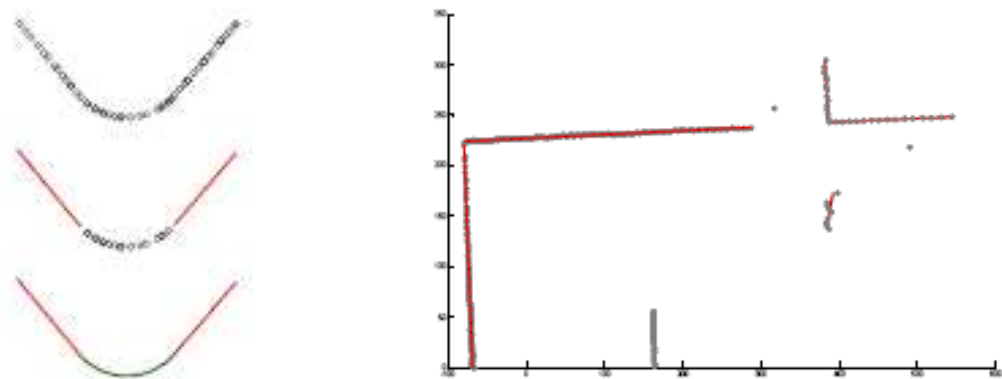


FIG. 2.12 – Résultat d'une détection d'arcs (à gauche) et de lignes (à droite) - figure extraite de [Xavier *et al.*, 2004].

## 2.3 L'extraction de primitives piétons

Cette section propose, dans un premier temps, d'identifier, lister et discuter des différents problèmes qui sont liés à la détection des piétons à l'aide d'un capteur laser dans un environnement routier urbain. Un état de l'art sur les différentes méthodes qui abordent explicitement le problème qui nous occupe ici est proposé. En effet, tous ces articles, revues ou thèses proposent des stratégies de détection, d'identification voire de suivi d'entités mobiles apparentées comme piétons à partir uniquement d'un capteur laser.

Dans un second temps, nous étudierons quelles sont les conditions qui sont requises et attendues afin d'obtenir un système de détection de piéton « idéal ». Ce travail permet de positionner les conditions de fonctionnement des algorithmes du projet LOVE vis-à-vis du détecteur « idéal ». En effet, devant la grande complexité du problème, le projet LOVE propose de réaliser un détecteur capable de détecter tous les piétons suivant certaines hypothèses.

### 2.3.1 Les problèmes de la détection des piétons

Pourquoi est-ce que la détection des piétons est si difficile à partir d'un capteur laser ? En effet, si pour un être humain, il est aisé d'accepter la définition suivante pour un piéton : « personne qui circule à pied », une telle assimilation n'est pas aisée pour un système de perception artificielle tel que celui examiné ici où les problèmes rencontrés tendent à s'accumuler.

#### Détection à bord d'un véhicule routier

Du fait que l'algorithme de perception soit monté sur un véhicule routier ajoute une difficulté supplémentaire surtout dans un environnement aussi complexe et varié qu'une ville. Effectivement, il est aisé d'imaginer la complexité de devoir détecter des piétons dans un environnement qui change tout le temps, à des vitesses différentes et où de nombreux obstacles peuvent apparaître et disparaître subitement à tout moment.

#### Détecter tous les piétons présents dans le champs du capteur

Le fait que tous les piétons, sans exception, doivent être détectés qu'ils soient mobiles ou pas, augmente encore la complexité de l'algorithme attendu. Evidemment, il est plus facile de se contenter de détecter voire de suivre des piétons uniquement en mouvement afin d'éviter toutes les fausses alarmes dues aux différents obstacles qui peuvent être rencontrés en ville tels que des panneaux, des barrières, des arbres et tous les objets qui peuvent s'apparenter par leur taille et/ou leur forme à un piéton.

#### Détecter malgré leurs postures ou leurs attitudes

Le piéton fait partie de ces objets les plus déformables qui existent avec un nombre insensé de poses, de tenues vestimentaires, de comportements et d'objets qui peuvent l'accompagner lors de ces déplacements. En effet, détecter un piéton qu'il soit adulte ou enfant, de face ou de côté, portant avec lui une valise ou un parapluie est totalement différent au niveau de sa taille et de sa forme dans l'image.



### Détecter même en cas d'occultation

Parce que le système doit travailler dans un trafic réel, un piéton peut à tout instant être soit partiellement soit totalement occulté par un autre objet. Le système doit être capable de suivre ces piétons pendant un certain temps malgré cette occultation partielle ou totale afin de pouvoir prévenir si une collision dangereuse est envisageable.

### Détecter rapidement

Evidemment, le système doit être en temps réel afin de prévenir d'une éventuelle collision. En effet, s'il ne peut pas détecter un piéton dans un temps suffisamment court, il est inutile. C'est une des principales complexités de la mise au point de ce genre d'algorithme qui doit pouvoir répondre à tous les problèmes énumérés jusqu'à présent tout en étant suffisamment simple pour un temps de calcul minimum.



FIG. 2.13 – A gauche, situation d'une apparition brutale d'un enfant - figure extraite de [Gavrila, 2001] . A droite, passage piéton avec de nombreuses occultations entre les usagers.



FIG. 2.14 – Exemples de différentes postures que peut prendre un piéton - figure extraite de [Gavrila *et al.*, 2001].

### Difficultés supplémentaires apportées par un capteur laser

Développer un algorithme de détection de piéton à partir d'un capteur laser comme celui choisi par le projet LOVe ajoute des difficultés supplémentaires qui sont liées aux



caractéristiques même des informations délivrées par le télémètre. On rappelle que les principales caractéristiques et surtout les limites d'utilisation de l'IBEO Alasca XT sont présentées dans la section 2.1.4.

Commençons par souligner quelques avantages liés au choix d'un tel télémètre. D'abord la mesure de distance (profondeur) est directement accessible sans traitement supplémentaire. Ce télémètre permet de capter instantanément quatre faisceaux laser et de percevoir les cibles suivant quatre coupes différentes. De plus, l'étendue de mesure (résolution de  $0.25^\circ$ ) est adaptée à l'application de perception d'un environnement routier immédiat ( $< 30$  mètres).

Par contre, comme pour tout capteur, les inconvénients ne manquent pas. Ici, de part le choix technologique effectué par les partenaires industriels avec un nombre de nappes et une résolution limitée, le capteur laser IBEO délivre une information pauvre des objets détectés (pas d'information de forme, de hauteur...), rendant de ce fait, plus difficile la classification des obstacles (cf. la figure 2.15).

De plus, sa forte sensibilité aux conditions atmosphériques complique la détection des obstacles qui peuvent être partiellement ou totalement occultés par le brouillard, la pluie ou la grêle. Le faisceau laser peut également être absorbé par certaines textures ou couleurs sombres des vêtements portés par les piétons rendant leur détection comme partiellement occultée.

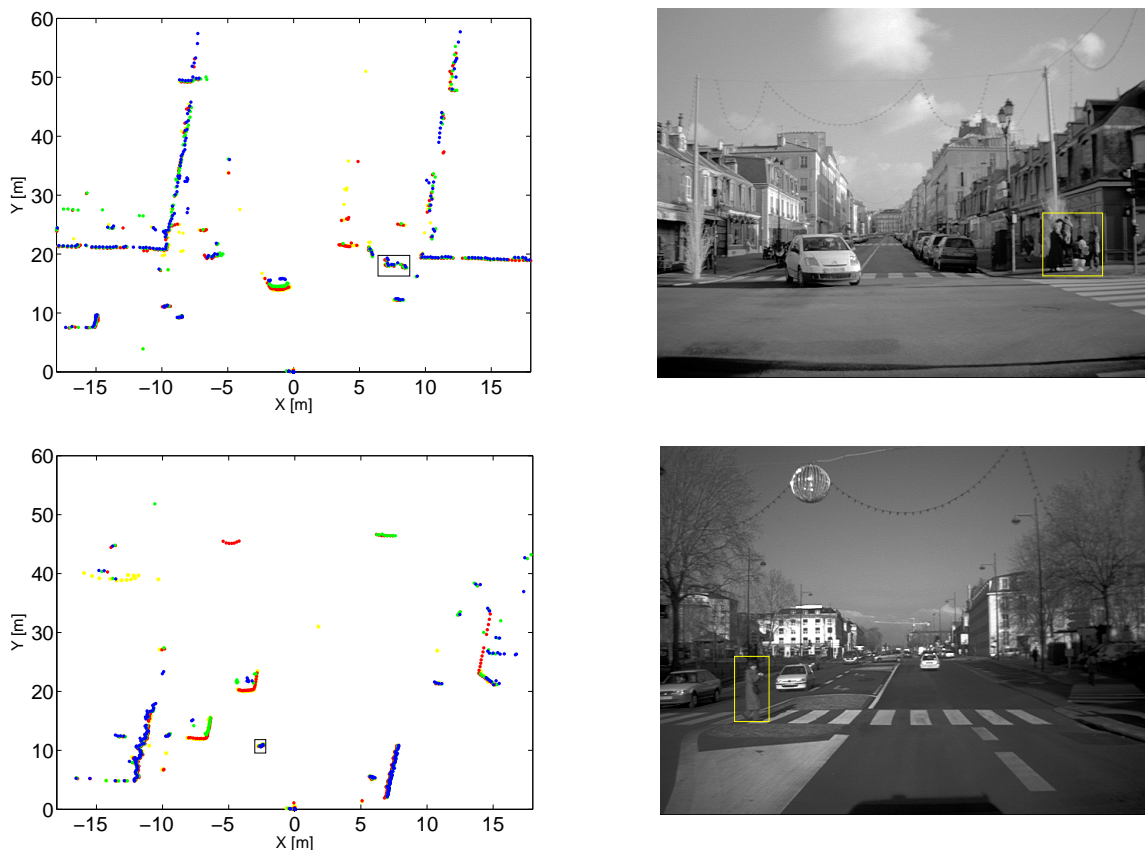


FIG. 2.15 – Ces deux images illustrent pourquoi la détection puis la classification des objets sont compliquées avec un capteur laser IBEO alasca XT. En effet, l'environnement perçu par ce capteur est « pauvre » c'est à dire que nous n'avons pas d'information sur la couleur, la hauteur ou la forme des objets.

Ces différents inconvénients peuvent être vus de manière générale comme les principaux problèmes d'une détection d'obstacles à partir d'une image de profondeur. En effet, un objet étant représenté dans une nappe, par un certain nombre de points 2D qui sont fonction de la distance objet/capteur et de la résolution angulaire du télémètre (cf. la figure 2.16), la plupart des méthodes de détection/classification vont tenter de classifier une observation (cluster) en fonction de sa taille et du nombre de points théoriques qui doit la constituer.

Si pour une raison quelconque (couleur absorbante, brouillard, transparence, occultation...) certains de ces points ne sont pas présents dans l'image laser, cela engendre une modification de la taille et/ou de la forme de l'obstacle qui risque de changer de catégorie lors de sa classification (cf. la figure 2.17).

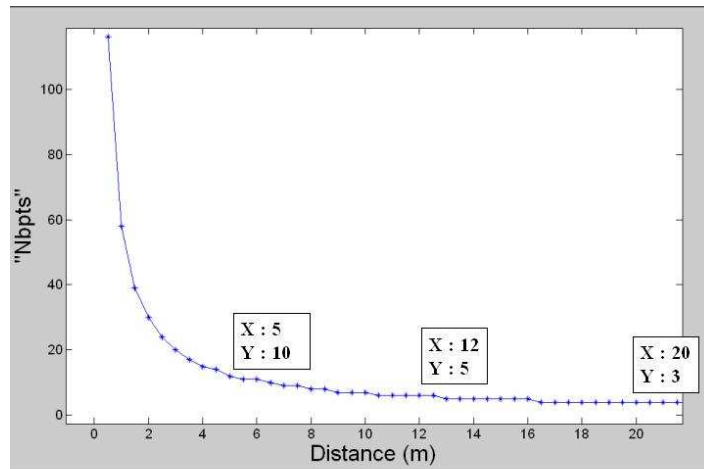


FIG. 2.16 – Nombre de points théoriques (pour une nappe) renvoyés par un piéton de profil (largeur  $\sim 40$  cm) pour un angle de résolution égal à  $0.5^\circ$  en fonction de la distance séparant le piéton du capteur.

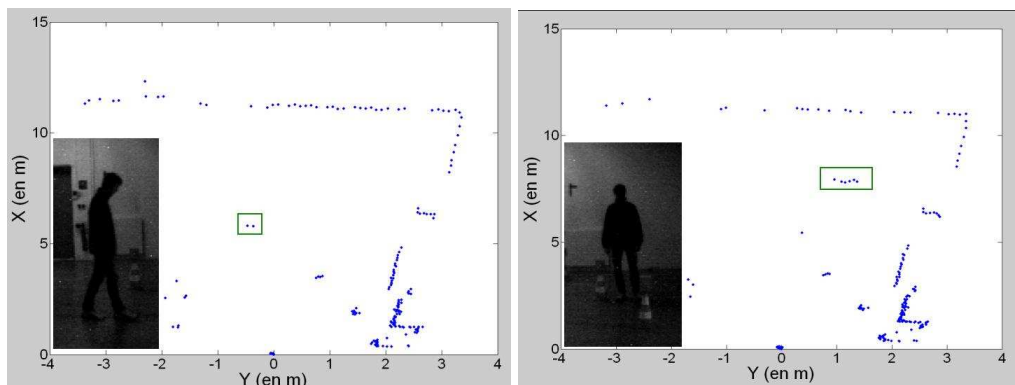


FIG. 2.17 – Le nombre de points renvoyés par un piéton dans une image laser peut être très différent selon sa posture et sa position relative par rapport au capteur. Par exemple, le piéton de gauche étant plus proche du capteur de détection que le piéton de droite, le capteur devrait renvoyer plus de points. Mais comme le piéton est de profil (surface de réflexion réduite), il renvoie moins de points que le piéton de face.

### 2.3.2 Etat de l'art des méthodes existantes

Dans cette section, est proposée une synthèse des différentes méthodes qui ont été développées ces dernières années sur la détection et/ou le suivi de piéton à partir de données issues uniquement d'un capteur laser. Durant ce travail, nous avons recensé 16 méthodes dans 16 articles, revues ou thèses depuis 1999 (cf. la figure 2.18). Ces méthodes sont regroupées dans deux contextes principaux :

- ▷ contexte statique : méthodes où le capteur laser est fixe,
- ▷ contexte dynamique : méthode où le capteur laser est mobile.

Ces deux catégories peuvent elles-mêmes être divisées en deux sous-catégories :

- ▷ les méthodes qui détectent et/ou suivent des piétons seulement en mouvement,
- ▷ les méthodes qui détectent et/ou suivent des piétons qui sont en mouvement ou non.

Par la suite, toutes les méthodes qui ont été retenues dans la figure 2.18 sont sommairement explicitées en les regroupant par catégorie de détections/classification utilisées.

#### Méthodes fondées sur les grilles d'occupation

Depuis une quinzaine d'années, la grille d'occupation présentée par Elfes [Elfes, 1989] est un outil très utilisé pour la construction de cartes en robotique mobile. La grille d'occupation est un découpage de l'espace dans lequel évolue le robot mobile en un ensemble de cellules. On cherche à estimer la probabilité qu'un objet (un piéton par exemple) occupe chacune de ces cellules à partir de mesures fournies par un capteur (laser par exemple). Cette estimation se fait incrémentalement, en considérant une à une les différentes observations.

Prassler [Prassler *et al.*, 1999] reconstruit les caractéristiques dynamiques des objets qui entourent un fauteuil roulant (cf. la figure 2.19) à partir de la différence de grilles d'occupation correspondant à des instants successifs.

Schulz [Schulz *et al.*, 2001] commence par extraire des objets en traitant chaque minimum local comme un amer qui représente un objet. Puis les entités mobiles ou statiques sont distinguées à l'aide d'une grille d'occupation (cf. la figure 2.20). Enfin, l'association de données et le suivi exploitent les filtres à particules sur une forme échantillonnée du filtre JPDAF.

Lindström [Lindstrom & Eklundh, 2001] utilise une méthode quelque peu différente permettant d'ignorer le problème de résolution et de taille de la grille d'occupation car sa technique ne divise pas la scène observée en cellules. En effet, il recherche dans les scans successifs des points qui « violent » (*violation points*) l'hypothèse d'environnement statique (carte 2D) (cf. la figure 2.21). Ces points sont modélisés par une impulsion gaussienne qui génère des hypothèses d'objets mobiles propagées à l'aide d'une technique de gradients.

**Commentaires :** Même si l'originalité de l'utilisation des grilles d'occupation tient au fait qu'elles ne prennent aucune décision intermédiaire (pas d'appariement piste/observation), cette méthode souffre de l'abstraction faite des objets qui sont détectés (ensemble de cellules) et rend difficile une classification robuste en tant que piéton. De plus, cette méthode permet principalement de ne détecter robustement que des objets en mouvement dans l'environnement ce qui limite le domaine de détection des piétons. Enfin, cette méthode fournit une représentation de l'environnement extrêmement riche ce qui demande un temps de calcul très important.

		niveau de difficulté																			
Lieu		intérieur										intérieur et/ou extérieur									
Véhicule	Obstacle	Statique				Statique et/ou en mouvement				Statique				Statique et/ou en mouvement							
		Mobile		Mobile et/ou statique		Mobile		Mobile et/ou statique		Mobile		Mobile et/ou statique		Mobile		Mobile et/ou statique					
Classification piéton ?	Nb plan(s) du capteur	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui				
hypothèses expérimentales																					
	1																				
	>1																				
	1		[Fod et al., 2002]																		
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	>1																				
	1																				
	&																				



FIG. 2.19 – Image du fauteuil roulant de Prassler circulant dans une station de métro - figure extraite de [Prassler *et al.*, 1999].



FIG. 2.20 – Suivi de personnes en intérieur avec à gauche une image de la séquence et à droite une visualisation 3D de la grille d'occupation - figure extraite de [Schulz *et al.*, 2001].

### Méthodes fondées sur le suivi du mouvement des jambes

Une des caractéristiques typiques permettant de reconnaître un piéton en mouvement est le déplacement périodique de ses jambes. En effet, il est possible de détecter le déplacement d'un piéton en se focalisant sur le mouvement périodique de ses deux jambes en utilisant le modèle de marche d'un individu proposé par Mochon et McMahon en 1980 [Mochon & McMahon, 1980] (cf. la figure 2.22).

Fuerstenberg [Fuerstenberg & Dietmayer, 2004] sont les premiers à proposer d'utiliser le mouvement périodique du déplacement des jambes d'un piéton dans une image laser (cf. les figures 2.23 et 2.24). En effet, les auteurs supposent que le mouvement de la jambe d'un piéton est harmonique c'est-à-dire qu'il décrit une fonction sinusoïdale. Par conséquent, le mouvement de la jambe est modélisé comme un pendule, où la position  $x^{Ped}(t)$  est caractérisée par l'équation différentielle :

$$\ddot{x}^{Ped}(t) = -w^2 x^{Ped}(t) \quad (2.2)$$

laquelle peut être résolue par l'équation :

$$x^{Ped}(t) = x_0^{Ped} \sin(wt) \quad (2.3)$$

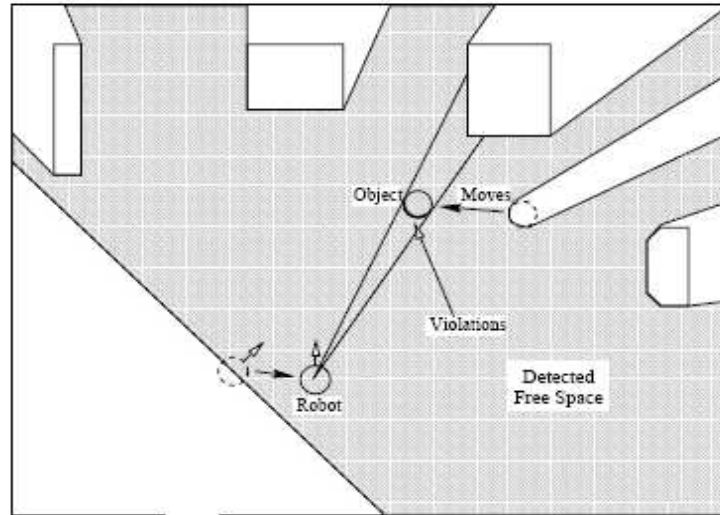


FIG. 2.21 – Exemple d’une grille d’occupation utilisée par Lindstöm pour détecter des obstacles en mouvement à partir de la violation des points - figure extraite de [Lindstrom & Eklundh, 2001].

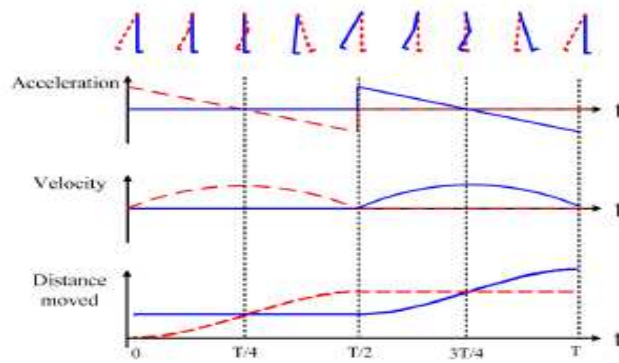


FIG. 2.22 – Modèle de marche d’un piéton - figure extraite de [Fuerstenberg & Dietmayer, 2004].

avec  $x_0^{Ped}(t)$  l’amplitude du déplacement, et  $w$  la fréquence angulaire du mouvement.

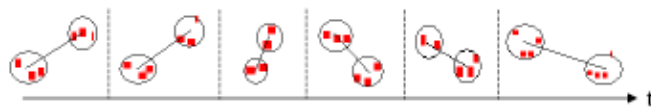


FIG. 2.23 – Détection des jambes d’un piéton dans une image laser à différents instants - figure extraite de [Fuerstenberg & Dietmayer, 2004].

Shao [Shao *et al.*, 2007], Zhao [Zhao & Shibasaki, 2005] et Cui [Cui *et al.*, 2007] ont des travaux similaires pour classer un objet comme piéton ou non. Tous utilisent le modèle de marche d’un humain pour décrire la périodicité du mouvement des pieds. Cette méthode leur permet de repérer les clusters susceptibles d’appartenir à un piéton (cf. la figure 2.24) à partir d’un ou plusieurs capteurs laser placés statiquement au sol (cf. la figure 2.25). Ils utilisent notamment les méthodes non paramétriques à noyau (KDE) pour



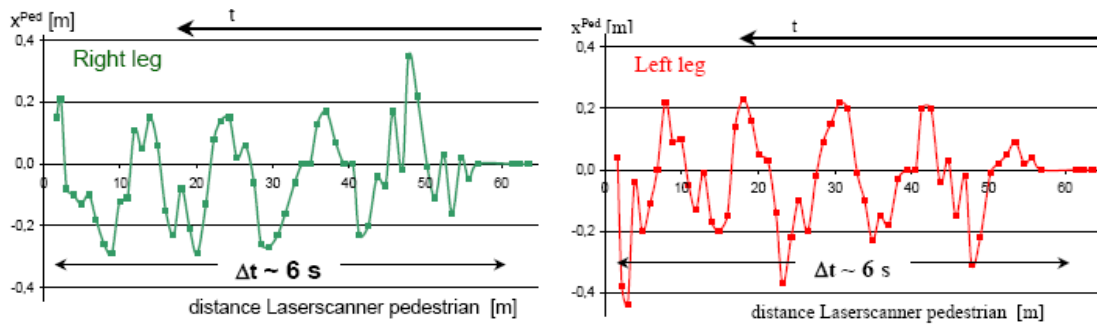


FIG. 2.24 – Enregistrement du mouvement de la jambe gauche (en rouge) et de la jambe droite (en vert) d'un piéton - figure extraite de [Fuerstenberg & Dietmayer, 2004].

classifier les différents clusters comme appartenant à des jambes de passant ou non. Shao et Zhao ajoutent pour trouver le maximum de vraisemblance une méthode d'optimisation déterministe fondée sur le principe du « mean-shift » [Grenier, 2005].

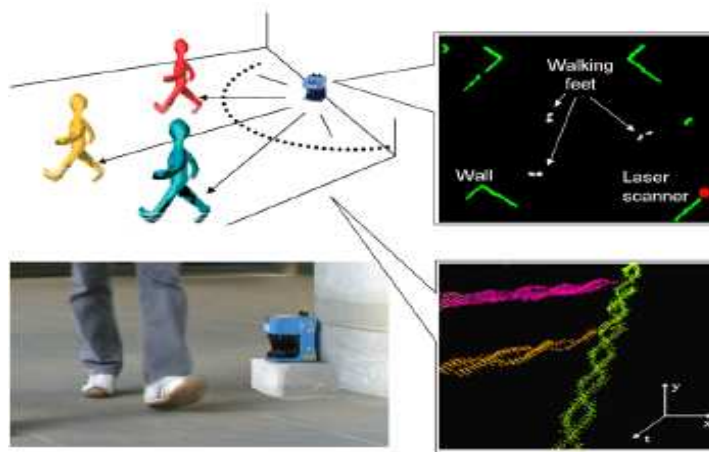


FIG. 2.25 – Le système de détection de piéton de Shao à partir d'un capteur laser. Dans l'image en bas à droite, nous pouvons remarquer le mouvement périodique des jambes qui a été détecté puis suivi - figures extraites de [Shao *et al.*, 2007]

**Commentaires :** Même si l'originalité de l'utilisation du modèle de marche d'un humain permet une détection robuste des piétons en mouvement, cette méthode souffre de l'impossibilité de détecter des piétons qui sont immobiles. De plus, cette méthode se réfère à un modèle de marche d'un humain, c'est-à-dire à une vitesse de déplacement d'environ 5 km/h. Donc si un piéton se déplace vite (pratique un jogging par exemple), l'algorithme arrive difficilement à le suivre. Enfin, cette méthode est mise à défaut dès qu'un piéton circule avec un bagage à la main, ce qui occulte le mouvement de ses jambes et rend sa détection difficile.

### Méthodes fondées sur une classification heuristique

Une classification heuristique est une méthode de classification qui consiste à mettre en relation des données relatives à un problème avec des solutions préétablies dans une base de connaissances. Toutes les méthodes décrites ci-dessous sont précédées d'une méthode de clustering permettant de regrouper les points laser voisins entre eux.

Afin de diminuer le nombre de fausses alarmes, Broggi [Broggi *et al.*, 2008] concentre l'attention de son algorithme de détection de piéton sur les zones dites « à risques » comme les coins de véhicules qui sont garées au bord de la route ou derrière les arrêts de bus (cf. la figure 2.26). Sa classification de cluster est simple : tous les petits clusters situés près des endroits à risques sont considérés comme piéton surtout s'ils sont en mouvement.

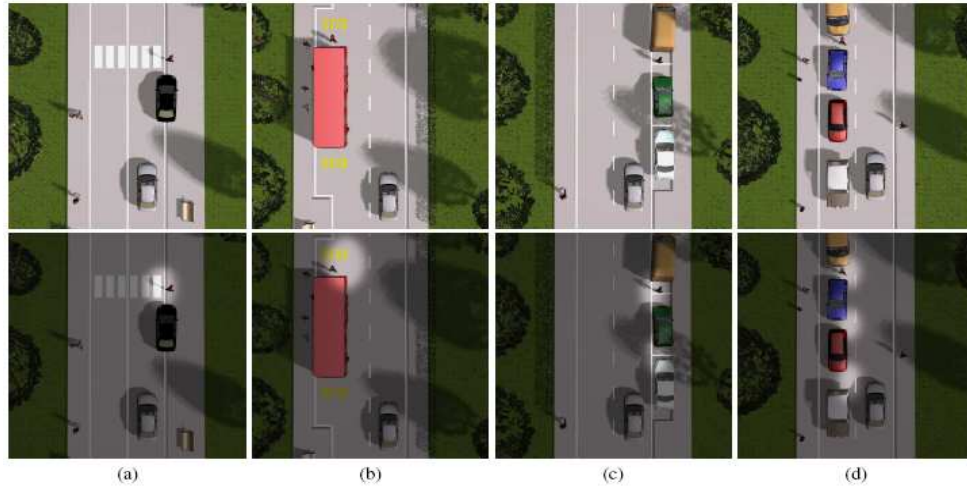


FIG. 2.26 – Diverses situations considérées comme dangereuses par les travaux de Broggi. Par exemple en (a), au niveau d'un passage piéton, un piéton est caché par un véhicule garé ; en (b), un autre est en train de traverser la route derrière un arrêt de bus ; en (c) et (d), un troisième traverse la route en passant entre deux véhicules garés au bord de la route - figures extraites de [Broggi *et al.*, 2008].

Dans les travaux de Navarro-Serment [Navarro-Serment *et al.*, 2008], tous les objets détectés sont testés afin de déterminer s'il s'agit de piétons. Dans ce but, chaque cible est évaluée par le calcul de son seuil de détection. Ce dernier mesure la confiance que l'algorithme accorde dans la classification d'un objet comme un piéton détecté. Ce seuil prend notamment en considération la taille théorique d'un piéton, sa vitesse de déplacement, et sa variation possible de taille suivant sa posture (cf. la figure 2.27).

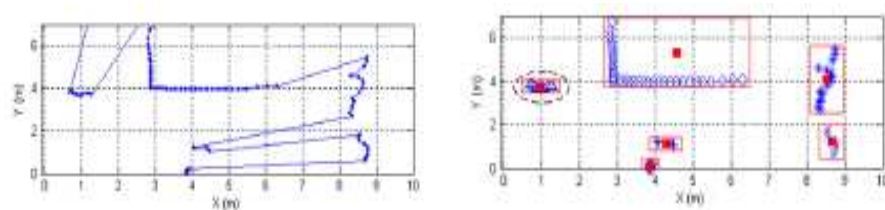


FIG. 2.27 – Illustration de la segmentation en ligne de l'image laser (à gauche), et classification (à droite) dans des boites englobantes des objets détectés dans l'image à partir du calcul du seuil - figures extraites de [Navarro-Serment *et al.*, 2008].

Fayad [Fayad *et al.*, 2008] s'appuie sur les caractéristiques intrinsèques du capteur laser utilisé afin de déterminer par une note de détection et de reconnaissance si les objets détectés sont effectivement ou non des piétons. Sa note de détection repose sur une connaissance *a priori* du nombre de points que doit renvoyer un piéton dans une image



laser en fonction de sa largeur (L) (2.4), de sa distance (D) au capteur et de l'angle de résolution ( $\alpha$ ) du capteur laser. Sa note de reconnaissance utilise une fonction probabiliste qui repose sur la largeur *a priori* du piéton en fonction de sa posture (cf. la figure 2.28).

$$N = \text{Round}\left(\frac{180 * 2}{\pi \alpha} \arctan\left(\frac{L}{2D}\right)\right) \quad (2.4)$$

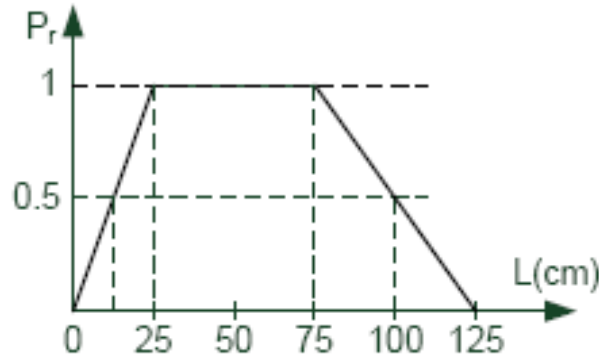


FIG. 2.28 – Fonction permettant de calculer la note de reconnaissance en fonction de la largeur du piéton - figure extraite de [Fayad *et al.*, 2008].

Fod [Fod *et al.*, 2002] soustrait un modèle de fond à l'image courante afin de ne conserver que les impacts situés au premier plan (cf. la figure 2.29). Ces impacts sont regroupés en « blobs » ou clusters qui sont mis en correspondance avec les blobs de l'image laser précédente afin de mettre à jour le nombre d'obstacles piétons détectés.

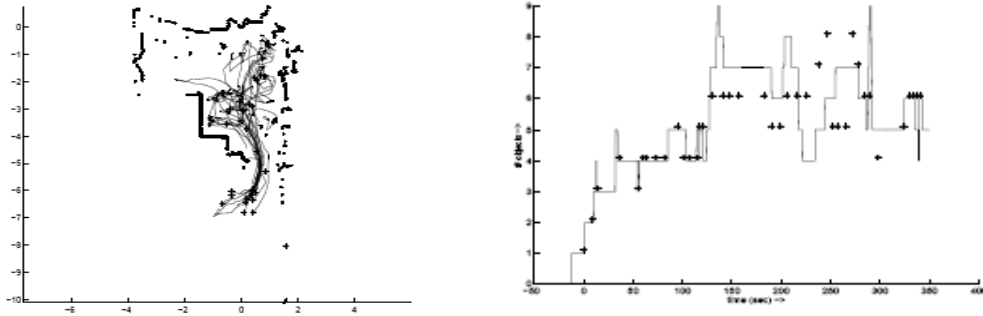


FIG. 2.29 – A gauche, exemple d'une détection et d'un suivi de piéton en vue de dessus dans un environnement intérieur. A droite, résultats temporels du suivi des objets dans cette pièce. Les étoiles représentent la vérité terrain tandis que les lignes représentent le résultat de détection et de suivi de l'algorithme - figures extraites de [Fod *et al.*, 2002].

Montemerlo [Montemerlo *et al.*, 2002] propose d'approximer un piéton par l'intersection d'un plan laser avec un cylindre (cf. la figure 2.30) et de suivre ces entités par un filtre à particules avec un modèle d'évolution brownien qui est un modèle d'évolution aléatoire qui revient à une marche aléatoire permettant d'effectuer des sauts discrets entre deux positions.

Enfin, Mendes [Mendes *et al.*, 2004] commence par regrouper en clusters les points voisins en distance avant de classer ces clusters par une méthode dite de « Vote à la

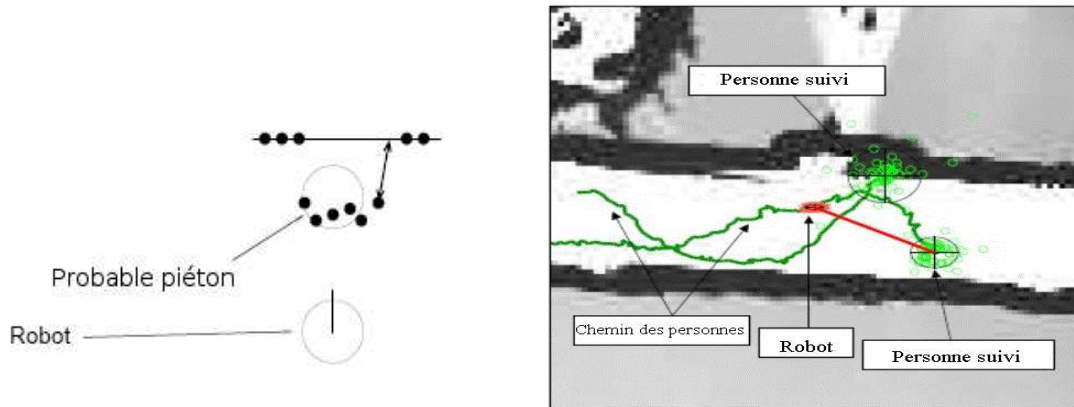


FIG. 2.30 – Illustration à gauche d’une détection de piéton en comparant un groupe de points avec un cylindre. A droite, exemple d’une détection et du suivi de deux piétons à partir du robot mobile - figures extraites de [Montemerlo *et al.*, 2002].

majorité » (cf. la figure 2.31) qui prend en compte un jeu de paramètres (acteur) tels que sa taille, le nombre de points, sa vitesse de déplacement... et conserve en tant que piétons tous ceux qui dépassent un certain seuil donné par le vote de chaque acteur.

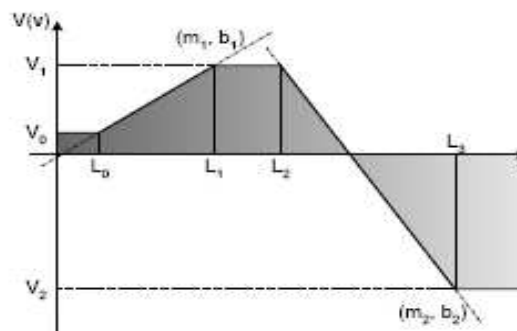


FIG. 2.31 – Modèle de la fonction de « Vote à la majorité » utilisée par [Mendes *et al.*, 2004].

**Commentaires :** Toutes les approches mentionnées ci-dessus, utilisent une classification heuristique fondée sur les caractéristiques géométriques des cibles détectées. En effet, sur des conditions de distance et de forme, ces algorithmes cherchent à distinguer dans l’image laser des clusters qui ont des formes de lignes, de cercles, d’amas de points et qui peuvent être regroupés afin de les classer comme des murs, des véhicules, des arbres, des barrières. Malheureusement si la classification heuristique est une stratégie bien adaptée et qui a fait ses preuves dans les cas de problème simple (véhicule, camion...), cette méthode reste encore largement défailante devant la classification d’objets aussi déformables que les piétons.

### Méthodes fondées SLAMMOT (Simultaneous Localization And Mapping and Moving Object Tracking)

L’objectif de la localisation et de la cartographie simultanée (SLAM) dans un environnement intégrant la détection et le suivi d’objets mobiles (MOT) est de maintenir une

carte de l'environnement d'un véhicule mobile, incluant des objets mouvants (cf. la figure 2.32). Le SLAMMOT permet notamment d'estimer conjointement la position du véhicule et les positions d'un ensemble d'objets de l'environnement (arbres, barrières, véhicules, piétons...) situé dans le champ de vue du capteur.

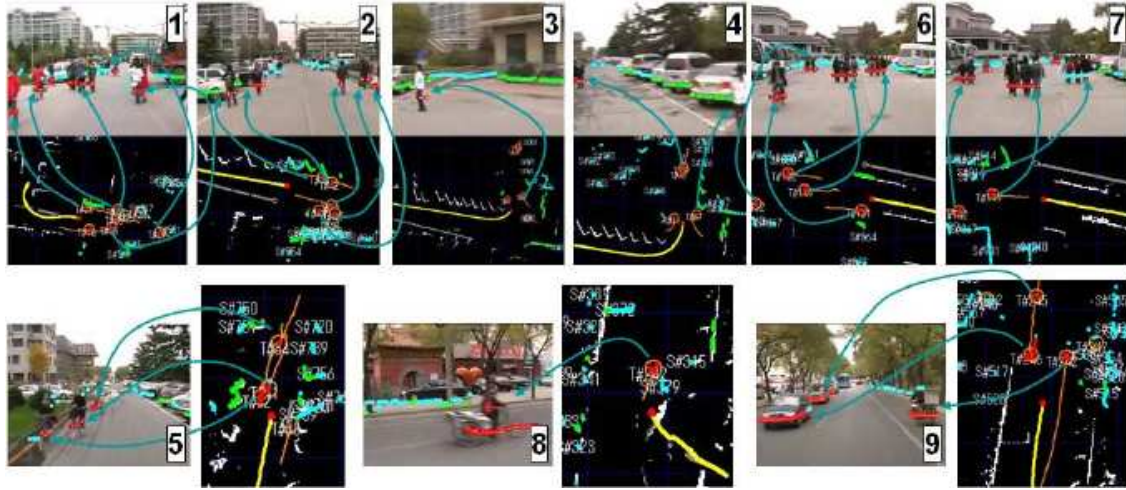


FIG. 2.32 – Résultats expérimentaux de SLAM avec des objets en mouvements - figures extraites de [Zhao *et al.*, 2008].

La majorité des algorithmes de classification d'objets fondés sur des capteurs laser, compte principalement sur les caractéristiques géométriques qu'ils peuvent trouver dans une image de profondeur. En effet, après avoir estimé leurs tailles, il suffit de les comparer à des boîtes englobantes prédéfinies à l'avance afin de les classifier comme tel ou tel objet. Malheureusement, dès que les objets sont mal détectés pour diverses raisons (conditions climatiques, occultation,...), la détection robuste de l'objet sur ce seul critère géométrique devient insuffisant, car déformé, il ne rentre plus dans sa boîte prédéfinie à l'avance, mais dans une autre et fausse la classification.

Afin de résoudre ce problème Gate [Gate & Nashashibi, 2008] et Zhao [Zhao *et al.*, 2008] proposent une approche originale similaire qui consiste à utiliser le SLAM [Wang, 2004] pour cartographier (enregistrer) l'environnement d'une image laser à une autre et de retrouver les objets précédemment détectés (cf. la figure 2.33). L'idée basique de cette méthode de détection de piéton fondée SLAMMOT est de considérer que la plupart des erreurs de classification vont être évitées si on est capable de retrouver (estimer) la vraie forme des cibles qui sont suivies malgré une mauvaise détection du capteur.

**Commentaires :** Cette méthode émet l'hypothèse que le problème de l'association de données entre les objets statiques et les objets dynamiques dans un environnement inconnu, large et complexe est complètement résolu. Malheureusement, ce n'est pas encore vrai ce qui peut mettre ce genre d'algorithme très vite en défaut. De plus, afin d'avoir une bonne estimation de la cartographie d'un environnement urbain aussi varié que complexe, l'algorithme a besoin d'un temps de calcul important ce qui est un problème dans le cadre d'une application automobile qui doit être temps réel.

### Autres méthodes utilisées pour classifier les piétons

D'autres auteurs ont testé des méthodes de classification qui ont donné satisfaction dans d'autres domaines comme dans la vision artificielle avec l'algorithme AdaBoost,

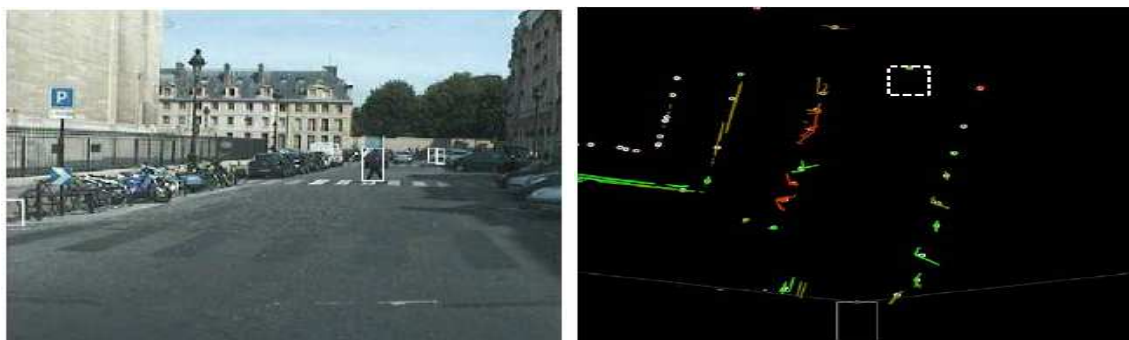


FIG. 2.33 – Détection de piéton à partir d'un capteur laser et de la méthode SLAMMOT. A gauche, une photo du lieu observé et à droite, les détections de piétons effectuées dans l'image laser - figures extraites de [Gate & Nashashibi, 2008].

l'informatique avec la théorie des graphes ou le tracking avec le modèle de mélanges gaussiens.

Kluge [Kluge *et al.*, 2001] commence par effectuer une segmentation heuristique de l'image laser par une agrégation de points en ensembles ou objets « quasi-convexes ». Puis il effectue un suivi de ces entités situées dans les images laser successives par une technique d'optimisation issue de la théorie des graphes. Un graphe (orienté) est une représentation symbolique simple permettant de modéliser toutes les relations possibles qui existent entre deux ensembles. Ce graphe permet donc à Kluge de prendre une décision sur l'association entre les objets situés dans deux images successives.

Monteiro [Monteiro *et al.*, 2006] propose de classifier les amas de points situés dans un scan laser par un classifieur GMM qui est préalablement paramétré par une étude des formes géométriques recherchées dans une image de profondeur. Les paramètres à définir étant comme pour toute fonction gaussienne, sa moyenne (taille de l'objet) et sa variance (variation possible de sa taille).

Enfin, Spinello [Spinello *et al.*, 2008] propose d'utiliser un algorithme de type AdaBoost pour classifier des clusters laser (cf. la figure 2.34). L'idée générale de l'algorithme Adaboost est que le résultat de plusieurs classifieurs est toujours meilleur que le résultat d'un seul. Adaboost permet donc de créer un classifieur dit « robuste » ou fort à partir de plusieurs classifieurs dit « faibles ». Spinello pense ainsi mieux classifier les clusters qui sont semblables pour entrer dans telle ou telle catégorie en demandant l'avis de plusieurs classifieurs.



FIG. 2.34 – Illustration d'une base de données comportant plusieurs « imagerie » de la trace d'un piéton dans une image d'impacts laser en vue de dessus utilisée par le classifieur Adaboost - figure extraite de [Spinello *et al.*, 2008].

**Commentaires :** Ces méthodes sont principalement des techniques d'apprentissage

qui permettent de classifier les différents clusters détectés. La principale motivation de l'exploration de ces différentes méthodes est d'essayer de se détacher d'une classification heuristique entre des clusters détectés dans une image laser et des formes géométriques choisies à l'avance. Cependant, ces méthodes ne permettent pas de résoudre le problème d'occultation des objets entre eux, ce qui pose toujours les mêmes problèmes de classification que pour les méthodes heuristiques.

### 2.3.3 De l'idéal à la réalité

Cette section commence par présenter quelques pistes pour répondre à la question : « Quelles sont les conditions requises d'un détecteur de piétons parfait ? ». Cette question se traite sur deux plans : le premier, évident, est l'expression formelle du cahier des charges d'un système parfait, le second, plus subtil, est celui de la technologie à employer pour y parvenir. Dans un contexte industriel réaliste immédiat, les objectifs sont plus modestes, mais doivent être réalisés avec rigueur dans un souci de validation permanente du cahier des charges. Pour y parvenir, le projet LOVE a inscrit l'ensemble de sa démarche dans le modèle du cycle en V qui est un modèle conceptuel de gestion de projet.

#### Un détecteur « idéal »

Le cahier des charges ci-dessous constitue l'idéal vers lequel un système de perception des vulnérables doit tendre (cf. la figure 2.35). Cependant, il semble aujourd'hui utopique d'espérer respecter tous ces points avec un unique capteur laser comme celui proposé par le projet LOVE.

- ▷ premièrement, le système doit être capable de détecter tous les piétons situés dans le champ du capteur, quels que soient leur comportement (en mouvement ou pas), leur attitude (de face, de côté,...), leur taille (grands, petits,...), leur nombre (seul ou en groupe) c'est-à-dire d'obtenir un taux de détection de 100%.
- ▷ deuxièmement, le système ne doit pas envoyer de fausses alarmes, quel que soit l'environnement traversé (parking, avenue, carrefour,...), c'est-à-dire d'obtenir un taux de fausses alarmes égal à 0.
- ▷ troisièmement, le système doit être capable de fonctionner par tous les temps (ensoleillé, nuageux, pluie, brouillard,...) et à toute heure, c'est-à-dire de nuit comme de jour.
- ▷ quatrièmement, le système doit être capable d'estimer précisément la position des piétons détectés puis de les identifier afin de les suivre, même en cas d'occultation partielle ou totale.
- ▷ enfin, le système doit être temps réel, c'est-à-dire que l'algorithme de détection doit être capable de traiter les données plus vite que la fréquence d'acquisition des capteurs.

Après lecture du cahier des charges d'un système de perception parfait, nous présentons deux capteurs qui permettraient de tendre plus facilement vers l'idéalisme de détection. Des appréciations sont portées sur les avantages et les inconvénients de ces deux principes de mesure en faisant remarquer dès à présent l'impossibilité industrielle d'utiliser aujourd'hui ce type de capteur encore trop onéreux.

#### Caméra thermique (ou infrarouge)

Une caméra thermique permet d'enregistrer les différents rayonnements infrarouges (ondes de chaleur) qui sont émis par les êtres vivants et qui varient en fonction de leur



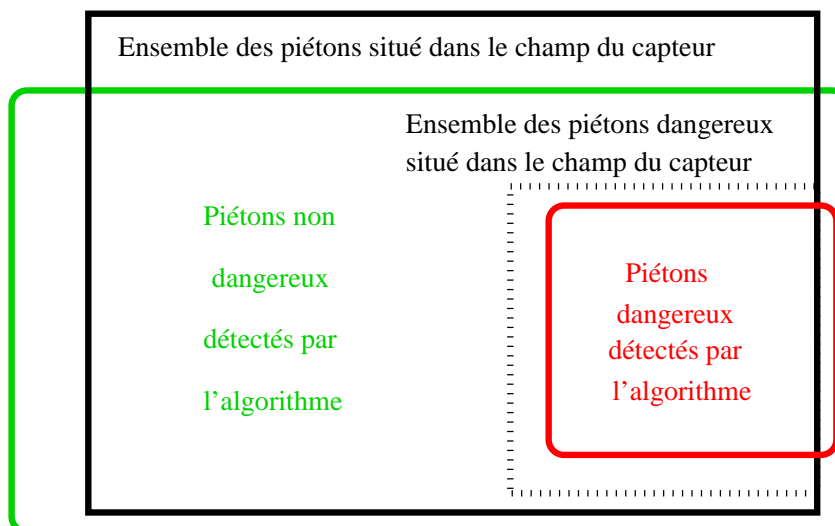


FIG. 2.35 – En noir, schéma d'un système de détection de piéton « idéal ». Un résultat probable d'algorithme de détection est représenté en vert avec son lot de fausses détections (contour en dehors du rectangle noir) et de non détection (contour à l'intérieur du rectangle noir). En pointillé, est représenté un système parfait de détection de piéton situé en situation de collision avec le véhicule alors qu'un résultat probable d'algorithme de détection est représenté en rouge avec également son lot de fausses détections et de non détection.

température. Ce capteur semble donc bien adapté pour détecter des corps chauds à température constante tel que ceux des piétons. De plus, ce capteur ajoute l'avantage de fonctionner de nuit comme par temps couvert (brouillard, neige, pluie...). Malheureusement, ces avantages disparaissent lors de journée chaude où les piétons et l'environnement ont un rayonnement infrarouge proche.

### Caméra 3D

En plus, des informations habituelles sur la couleur et la forme des objets observés, une caméra 3D permet d'ajouter l'information de profondeur (sur le principe de la stéréovision) qui fait cruellement défaut à une caméra 2D classique. La caméra 3D constitue certainement le capteur présentant le plus d'avantages par la richesse de l'information enregistrer. Cependant, ces derniers disparaissent par visibilité réduite (brouillard, pluie, soleil couchant...) et de nuit.

### Commentaires

Au vu des qualités de ces deux capteurs, nous pouvons légitimement penser qu'une combinaison entre une caméra thermique et une camera 3D semble être la meilleure solution garantissant l'obtention d'images utilisables dans toutes les conditions climatiques et à toute heure (cf. la figure 2.36).

### Un détecteur à hypothèses « réalistes »

Le projet LOVe a choisi d'étudier un système de détection des vulnérables pouvant fonctionner avec des capteurs de perception aux performances compatibles avec une exploitation industrielle rapide. Le présent cahier des charges tient donc compte des limites des capteurs utilisés et propose de réaliser un détecteur de piéton fiable et sûr dans des conditions d'utilisation restreintes vis-à-vis d'un système idéal.



FIG. 2.36 – Exemple typique d’une scène observée dans une caméra opérant dans le visible (à gauche) et dans l’infrarouge (à droite) - figures extraites de [Bertozzi *et al.*, 2007]

### Définition du contexte d’utilisation

Le système doit fonctionner de jour dans des environnements urbains complexes comme ceux rencontrés dans les grandes villes françaises. Par conséquent, la vitesse du véhicule ne dépasse pas les 50 *km/h*.

### Définition d’un piéton à détecter

On considère que le système doit détecter chaque piéton individuellement. Le terme « individuellement » signifie que chaque piéton dont le buste peut être isolé (dans le cadre d’une image) doit être identifié par le système (cf. la figure 2.37). Cela étant, si le système parvient à identifier et à suivre des piétons masqués (partiellement ou complètement) cette information sera considérée comme correcte.



FIG. 2.37 – En vert, un piéton qui doit être absolument détecté (buste visible) et en rouge, un piéton masqué dont la détection n’est pas requise *a priori*.

### Définition d’un piéton à risque

Les piétons à détecter absolument sont situés dans la zone à risque définie dans la figure 2.38. On considère qu’un piéton est en position d’être percuté par un véhicule s’il est situé à une distance inférieure à deux secondes de la vitesse du véhicule. Comme pour

la définition des piétons à détecter, si le système parvient à identifier et à suivre des piétons situés hors de la zone à risque, cette information sera considérée comme correcte.

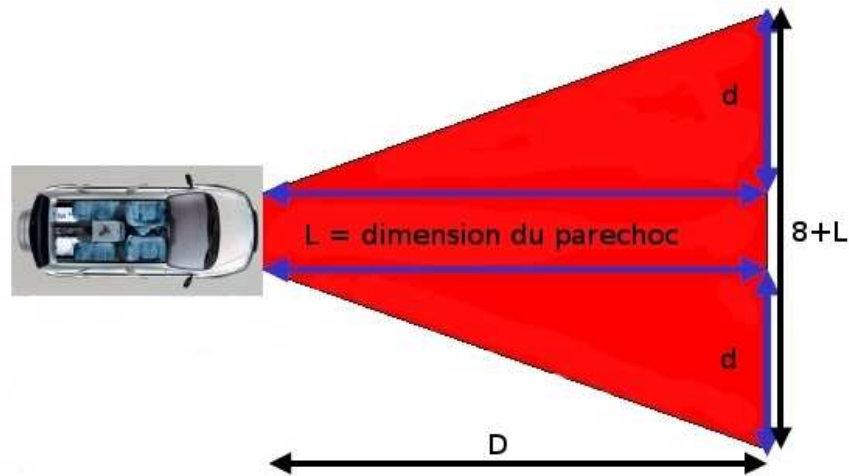


FIG. 2.38 – Un piéton est à risque si et seulement s’il est présent dans la zone rouge. Cette zone est définie de la façon suivante : on considère un piéton se déplaçant à  $2 \text{ m/s}$  perpendiculairement à la trajectoire du véhicule. La dimension du pare-chocs étant constante (par exemple  $L = 2 \text{ m}$ ), les piétons pouvant être heurtés sont à une distance  $V_{\text{piéton}} * t_{\text{avant\_impact}}$  du pare-chocs soit  $2 * 2 = 4 \text{ mètres}$ .

### Occultations temporaires

Tout piéton, présent dans le champ de vision du capteur devra être détecté et signalé. Cette détection devra être maintenue autant que possible en cas de masquage temporaire du piéton par un autre objet. En cas de masquage, la durée de maintien d’une piste sera au minimum de  $500 \text{ ms}$  car un piéton se déplaçant à  $2 \text{ m/s}$  parcourt  $1 \text{ mètre}$  en  $500 \text{ ms}$ . Un suivi de  $500 \text{ ms}$  devrait donc être suffisant pour gérer les croisements entre les piétons. Une valeur supérieure est envisageable pour gérer les occultations entre un piéton et un véhicule occultant dont la largeur est plus proche des  $2 \text{ mètres}$ .

Toutes les données et/ou détections de piétons seront exprimées dans le repère du capteur LiDAR ( $O, X, Y, Z$ ) avec  $O$  au centre du pare-chocs (emplacement du LiDAR),  $X$  vers l’avant,  $Z$  vers le haut et  $X$ - $Y$  le plan parallèle au sol.

## 2.4 Conclusion

Dans ce chapitre, nous avons décrit dans un premier temps, le principe de la télémétrie laser, en présentant les avantages et les inconvénients d’une détection d’objets à l’aide de données laser. Un état de l’art technologique sur les différents capteurs laser existant sur le marché a permis de situer les performances du capteur choisi dans le cadre du projet LOVE. Bien que performant en détection (notamment en précision), ce capteur souffre de désavantages importants. D’une part, la résolution angulaire ( $\geq 0.25^\circ$ ) ne permet plus d’obtenir un relevé télémétrique dense de l’environnement après  $30 \text{ m}$ . D’autre part, le nombre de nappes limité ( $4 \leq$ ) ne permet plus d’obtenir une « réelle » information 3D, car nous perdons l’information de hauteur des objets.



Ensuite, une description des problèmes qui sont spécifiques à la détection des piétons nous a permis de lister et d'identifier les principaux points scientifiques (environnement urbain, objets très déformables) et techniques (avantages et inconvénients du télémètre Alasca XT, capteur embarqué) qui sont à résoudre.

Par la suite, un état de l'art des différentes méthodes qui ont été développées ces dernières années sur la détection des piétons à partir d'un unique capteur laser a été proposé et discuté. Tous ces articles ont été regroupés par catégorie en rappelant de façon très schématique l'approche algorithmique afin de permettre une lecture rapide et claire de leur méthode. On peut remarquer que la plupart de ces dernières sont fondées sur l'utilisation d'un capteur laser mono-plan et que la plupart des auteurs se contentent de détecter les piétons en mouvement afin d'éviter des erreurs de reconnaissances entre les « objets-piétons » avec les autres obstacles de la route comme par exemple les panneaux de circulation. En revanche, il est souvent difficile d'évaluer la pertinence des algorithmes proposés car peu d'auteurs explicitent clairement les performances de détection ou non de leur méthode.

Enfin, ce chapitre se termine en listant les qualités d'un détecteur de piéton parfait permettant de situer le niveau d'exigence attendu par le cahier des charges du projet LOVE.

---

## Chapitre 3

# Le suivi d'objet à partir d'un ou plusieurs capteurs

Le suivi d'objet consiste à suivre dans le temps les positions des différentes « détections d'objets », à partir de mesures émises par un ou plusieurs capteurs. Cette fonctionnalité est généralement nécessaire pour pallier aux données souvent incomplètes et bruitées délivrées par les capteurs. En effet, en conservant une trace de chaque objet, il est possible d'enrichir notre connaissance de la scène observée afin d'améliorer la détection de ces objets. De plus, en fonction du besoin, le système de suivi peut ajouter de la connaissance aux objets détectés en prenant en charge l'estimation de la vitesse qui ne peut pas être directement mesurée par un capteur laser par exemple. Ce chapitre commence par présenter le principe du filtrage bayésien. Nous rappelons ensuite les différents algorithmes de filtrage bayésien, à savoir le filtre de Kalman et ses extensions, ainsi que les méthodes de Monte Carlo tels que les filtres SIR et MCMC (Markov Chain Monte Carlo). Chaque méthode est commentée en détaillant leurs avantages mais aussi leurs limites. Ce chapitre se termine sur l'extension du problème de suivi multi-objets au cas multi-capteurs en proposant dans un premier temps, un état de l'art des principales méthodes existantes avant de terminer sur les différentes architectures de fusion qui sont habituellement rencontrées.

### 3.1 Modélisation du problème de suivi d'objet

En suivi d'objet, comme en détection ou en classification (cf. la section 2 du chapitre 2), l'opération consiste à estimer l'état de la variable cachée  $\mathbf{X}$  c'est-à-dire « ce que l'on cherche » à partir de l'observation  $\mathbf{Z}$ , délivrée par un ou plusieurs capteurs. Ce problème de suivi peut-être résolu à travers deux stratégies [Bardet, 2009] illustrées par la figure 3.1.

- ▷ l'approche inverse,
- ▷ l'approche directe.

#### 3.1.1 Méthodes inverses

Ces stratégies consistent à construire l'état  $\mathbf{X}$  d'une scène réelle à partir de l'observation  $\mathbf{Z}$ , en appliquant la transformation  $\mathbf{X} = g(\mathbf{Z})$ . Elles fournissent alors directement une estimation de l'état à partir des observations, sans connaissance aucune de la scène observée. Ces approches sont couramment utilisées pour la détection d'obstacle, où l'état  $\mathbf{X}$  se réduit en une variable binaire définissant la présence ou non d'objet.

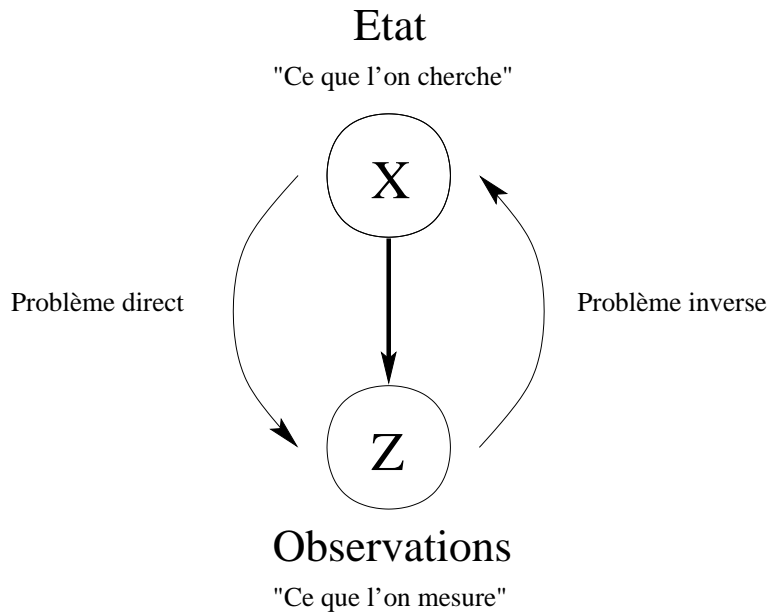


FIG. 3.1 – Schéma des deux approches permettant d'estimer l'état de la variable cachée qui représente notre connaissance de la scène. L'estimation de l'état à partir d'une observation est un problème inverse. Le processus qui consiste à construire une observation à partir de l'état est un problème direct.

La plupart de ces méthodes sont fondées sur l'apprentissage hors ligne, utilisant une base d'exemples créées à l'avance. Les algorithmes que l'on rencontre le plus souvent sont basés sur les machines à vecteur de support (SVM) ou le boosting [Arras *et al.*, 2007].

### 3.1.2 Méthodes directes

Ces stratégies consistent à construire une observation à partir d'un état  $X$  connu *a priori*. On ne cherche donc pas à expliciter la fonction de transformation inverse  $X = g(Z)$ , mais on cherche la transformation  $Z = f(X)$  qui permet de faire correspondre une observation avec une prédiction de l'état  $X$ . Nous sommes alors dans une approche basée modèle, qui demande une connaissance sur : le modèle de l'objet, le modèle de l'observation qu'il génère et une fonction de vraisemblance mesurant la concordance entre l'observation  $Z$  et l'état prédit  $X$ .

Il existe à ce jour, deux principales classes de stratégies directes qui sont les méthodes probabilistes et les méthodes déterministes. Les méthodes déterministes consistent à rechercher les paramètres de mouvement optimaux permettant de minimiser l'erreur entre un modèle de l'objet suivi et l'information extraite de chaque nouvelle observation  $Z$ . Elles se réduisent donc généralement à un problème d'optimisation [Baker & Matthews, 2004]. Les méthodes probabilistes quant à elles, ne se contentent pas d'estimer un état optimum *a posteriori* mais estiment toute la loi de probabilité *a posteriori* de cet état. L'information délivrée est donc plus riche.

**Notre algorithme de suivi repose sur une stratégie directe, fondée sur le filtrage bayésien qui cherche l'état caché de  $X$  en s'appuyant sur le théorème de Bayes.**

## 3.2 Formalisation du filtre bayésien

Le problème qui nous intéresse peut se formuler comme suit. On considère un système dynamique composé de deux vecteurs aléatoires à temps discret  $\mathbf{X}$  et  $\mathbf{Z}$ .

Le vecteur  $\mathbf{X}$  est appelé processus d'état. Il est inconnu et supposé markovien de distribution initiale  $p\{\mathbf{X}_0\}$  et de loi de transition  $p(\mathbf{X}_k|\mathbf{X}_{k-1})$ . Cette loi de transition décrit le modèle d'évolution de l'état.

Le vecteur  $\mathbf{Z}$  est appelé vecteur d'observation. Les observations  $\{Z_{1:k} = Z_1, \dots, Z_k\}$  sont supposées être indépendantes. Le vecteur  $\mathbf{Z}_k$  est supposé conditionnellement indépendant au processus d'état et sa distribution ne dépend que de l'état au même instant. Le processus d'observation sera donc décrit par la distribution  $p(\mathbf{Z}_k|\mathbf{X}_k)$ , aussi appelée vraisemblance. On appellera équation de mesure, l'équation associée qui relie l'observation à l'état.

La propriété markovienne du processus d'état ainsi que l'indépendance conditionnelle du processus d'observation amènent à modéliser le système dynamique par une chaîne de Markov cachée représentée par le graphe d'indépendance de la figure 3.2.

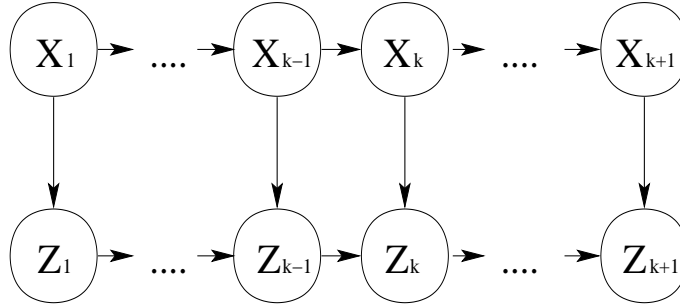


FIG. 3.2 – Modèle graphique du réseau bayésien dynamique pour un suivi multi-objets.

La résolution du problème de filtrage consiste alors à construire un algorithme itératif pour l'estimation de la distribution *a posteriori* de l'état à l'instant courant conditionnellement à l'ensemble des mesures passées et présentes  $p(\mathbf{X}_k|\mathbf{Z}_{1:k})$ . Le filtre bayésien optimal nous fournit une solution exacte en deux étapes.

- ▷ en supposant  $p(\mathbf{X}_{k-1}|\mathbf{Z}_{1:k-1})$  connue, l'étape de prédiction permet de connaître la distribution prédite. C'est l'hypothèse markovienne et l'équation de Chapman Kologorov qui permettent l'écriture de cette étape.

$$p(\mathbf{X}_k|\mathbf{Z}_{1:k-1}) = \int p(\mathbf{X}_k|\mathbf{X}_{k-1})p(\mathbf{X}_{k-1}|\mathbf{Z}_{1:k-1})d\mathbf{X}_{k-1} \quad (3.1)$$

- ▷ l'étape de correction introduit la nouvelle observation, permettant de mettre à jour la prédiction. Le théorème de Bayes nous permet alors, d'obtenir une expression de la distribution recherchée.

$$p(\mathbf{X}_k|\mathbf{Z}_{1:k}) = \frac{p(\mathbf{Z}_k|\mathbf{X}_k)p(\mathbf{X}_k|\mathbf{Z}_{1:k-1})}{p(\mathbf{Z}_k|\mathbf{Z}_{1:k-1})} \quad (3.2)$$

Notons que le terme  $p(\mathbf{Z}_k|\mathbf{Z}_{1:k-1})$  est constant et sera noté  $C$  car les observations sont indépendantes les unes des autres (aucun raccordement sur le graphe bayésien de la figure 3.2). Donc, en regroupant les équations 3.1 et 3.2, on obtient l'équation du filtre

bayésien récursif :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) = C^{-1} p(\mathbf{Z}_k | \mathbf{X}_k) \int p(\mathbf{X}_k | \mathbf{X}_{k-1}) p(\mathbf{X}_{k-1} | \mathbf{Z}_{1:k-1}) d\mathbf{X}_{k-1} \quad (3.3)$$

Cette équation donne la fonction de densité de probabilité *a posteriori* recherchée à partir des trois quantités suivantes :

- ▷  $p(\mathbf{Z}_k | \mathbf{X}_k)$  l'équation de vraisemblance. C'est l'adéquation entre une mesure (observation) et une hypothèse prédite de l'état courant des objets.
- ▷  $p(\mathbf{X}_k | \mathbf{X}_{k-1})$  le modèle d'évolution. Ce modèle doit permettre de prédire l'état courant  $\mathbf{X}_k$  à partir de l'état précédent  $\mathbf{X}_{k-1}$ . Ce modèle modélise la dynamique de l'état (position, vitesse...).
- ▷  $p(\mathbf{X}_{k-1} | \mathbf{Z}_{1:k-1})$  la distribution de probabilité de l'état précédent. Elle représente notre connaissance précédente de l'état des objets suivis dans la scène observée.

De part la très grande dimension des intégrales mises en jeu par le filtre bayésien optimal (3.3), un calcul direct de ces expressions n'est pas possible dans le cas général. Elle peut être calculée soit sous forme analytique (méthodes paramétriques telles que les distributions gaussiennes), soit sous forme approximée (méthodes stochastiques telles que les techniques de Monte Carlo). Les figures 3.3 et 3.4 illustrent ces deux possibilités.

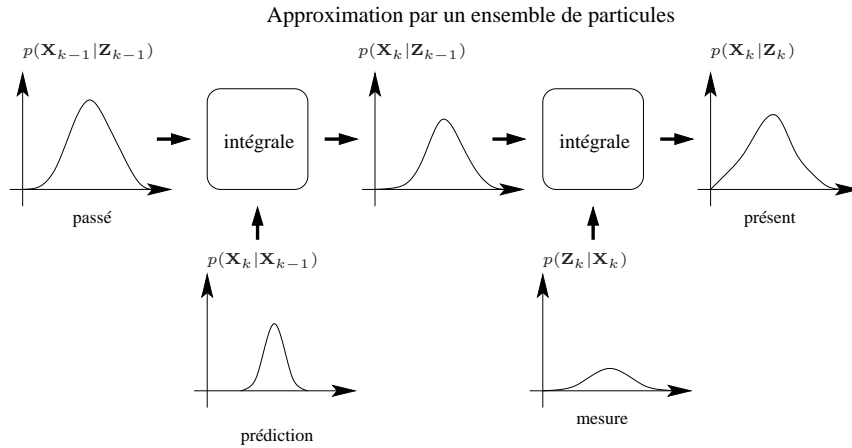


FIG. 3.3 – Représentation schématique du filtre bayésien récursif par les méthodes paramétriques.

### 3.3 Méthodes paramétriques

#### 3.3.1 Modèle linéaire avec bruit gaussien : filtre de Kalman

Le filtre de Kalman [Kalman, 1960] et son extension au temps continu appelée filtre de Kalman-Bucy [Kalman & Bucy, 1961] permettent de résoudre de façon optimale le problème de filtrage linéaire, quand les bruits du système sont additifs et gaussiens. Le principe de ce filtre est d'estimer l'état d'un système dynamique à partir d'une série incomplète de mesures bruitées. L'estimation de l'état courant du système (instant  $k$ ) se fait uniquement à partir de la mesure obtenue à l'instant précédent. Ainsi, on considère le

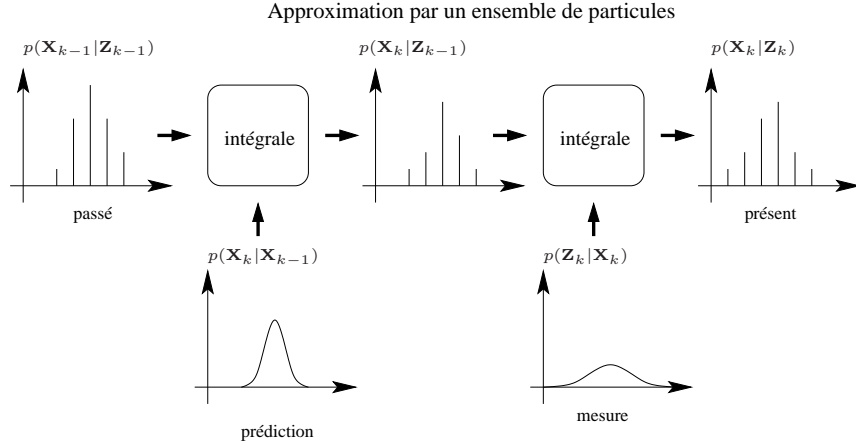


FIG. 3.4 – Représentation schématique du filtre bayésien récursif par les méthodes séquentielles de Monte Carlo.

système d'équation linéaire suivant :

$$\begin{aligned} \mathbf{X}_k &= \mathbf{F}_k \mathbf{X}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k \\ \mathbf{Z}_k &= \mathbf{H}_k \mathbf{X}_k + \mathbf{D}_k \mathbf{u}_k + \mathbf{v}_k \end{aligned} \quad (3.4)$$

où les hypothèses suivantes sont faites :

- ▷ l'état initial  $\mathbf{X}_0$  est gaussien, d'espérance  $\hat{\mathbf{X}}_0$  et de covariance  $\Sigma_0$ ,
- ▷ les matrices  $\mathbf{F}_k$  et  $\mathbf{H}_k$  sont déterministes, appelées respectivement matrices d'état et de mesure,  $\mathbf{B}_k \mathbf{u}_k$  et  $\mathbf{D}_k \mathbf{u}_k$  sont des entrées du système connues,
- ▷ les bruits d'état  $\mathbf{w}_k$  et de mesure  $\mathbf{v}_k$  sont des bruits blancs gaussiens de moyennes nulles et de covariances respectives  $\mathbf{Q}_k$  et  $\mathbf{R}_k$  connues. Ils sont supposés mutuellement indépendants et indépendants de la condition initiale.

Le système peut alors s'écrire de la façon suivante :

$$\begin{aligned} \mathbf{X}_0 &\sim \mathcal{N}(\mathbf{X}_0; \hat{\mathbf{X}}_0, \Sigma_0) \\ \mathbf{X}_k | \mathbf{X}_{k-1} &\sim \mathcal{N}(\mathbf{X}_k; \mathbf{F}_k \mathbf{X}_{k-1} + \mathbf{B}_k \mathbf{u}_k, \mathbf{Q}_k) \\ \mathbf{Z}_k | \mathbf{X}_k &\sim \mathcal{N}(\mathbf{Z}_k; \mathbf{H}_k \mathbf{X}_k + \mathbf{D}_k \mathbf{u}_k, \mathbf{R}_k) \end{aligned} \quad (3.5)$$

La notation  $\mathcal{N}(\mathbf{X}; \hat{\mathbf{X}}, \Sigma)$  désigne la loi normale de variable  $\mathbf{X}$ , d'espérance  $\hat{\mathbf{X}}$  et de covariance  $\Sigma$ . Par linéarité des équations d'état et de mesure, le processus  $\{\mathbf{X}_k, \mathbf{Z}_k\}$  est gaussien et la loi de filtrage recherchée  $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$  est gaussienne. La loi de filtrage est donc entièrement décrite par son espérance  $\hat{\mathbf{X}}_{k|k} = \mathbb{E}[\mathbf{X}_k | \mathbf{Z}_{1:k}]$  et sa covariance  $\Sigma_{k|k} = \mathbb{E}[(\mathbf{X}_k - \hat{\mathbf{X}}_{k|k})(\mathbf{X}_k - \hat{\mathbf{X}}_{k|k})^t]$ . Ces deux premiers moments sont calculés récursivement par le filtre de Kalman en deux étapes :

1. l'étape de **prédiction** qui consiste à estimer l'état courant à partir de l'estimation de l'état précédent. Pendant cette phase on calcule :
  - ▷ une estimation de l'état courant  $\hat{\mathbf{X}}_{k|k-1}$ ,
  - ▷ la matrice de covariance de l'erreur d'estimation  $\hat{\Sigma}_{k|k-1}$ .
2. l'étape de **correction** (ou mise à jour de mesure) qui utilise les observations de l'état courant pour corriger sa prédiction courante et ainsi augmenter la précision de l'estimation. Pendant cette phase on calcule :
  - ▷ la matrice de covariance de l'innovation  $\mathbf{S}_k$ , qui indique l'écart entre la mesure observée et sa prédiction,

- ▷ le gain du filtre de Kalman  $\mathbf{G}_k$ , doit tenir compte des incertitudes relatives de l'estimation courante et de la donnée. Il faut noter que ce dernier devient d'autant plus important que la covariance du bruit de la mesure  $\mathbf{R}_k$  est négligeable devant celle du modèle  $\hat{\Sigma}_{k|k-1}$ . A l'inverse, si l'incertitude de la donnée est grande par rapport à celle de l'estimation, le gain sera faible (donnée peu fiable), il est donc normal que le gain modifie peu l'estimation courante,
- ▷ l'estimation de la mise à jour de l'état  $\hat{\mathbf{X}}_{k|k}$  et de la matrice de covariance  $\hat{\Sigma}_{k|k}$ .

Ces deux étapes sont résumées dans l'algorithme 1.

---

**Algorithme 1** Filtre de Kalman

---

- **Initialisation :**  $\hat{\mathbf{X}}_{0|0} = \hat{\mathbf{X}}_0, \Sigma_{0|0} = \Sigma_0$   
pour  $k = 1, 2, \dots, K$ .
  - **Prédiction :**

$$\hat{\mathbf{X}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{X}}_{k-1|k-1} + \mathbf{B}_k \mathbf{u}_k$$

$$\Sigma_{k|k-1} = \mathbf{F}_k \Sigma_{k-1|k-1} \mathbf{F}_k^t + \mathbf{Q}_k$$
  - **Correction :**

$$\mathbf{S}_k = \mathbf{H}_k \hat{\Sigma}_{k|k-1} \mathbf{H}_k^t + \mathbf{R}_k$$

$$\mathbf{G}_k = \hat{\Sigma}_{k|k-1} \mathbf{H}_k^t \mathbf{S}_k^{-1}$$

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{G}_k [\mathbf{Z}_k - (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1} + \mathbf{D}_k \mathbf{u}_k)]$$

$$\hat{\Sigma}_{k|k} = [\mathbf{I} - \mathbf{G}_k \mathbf{H}_k] \hat{\Sigma}_{k|k-1}$$
- 

**Commentaires**

Dans le cas des systèmes linéaires et gaussiens où l'évolution de la loi conditionnelle est déterminée par sa moyenne et sa matrice de covariance, le filtre de Kalman fournit la solution optimale sous la forme d'une expression récursive des deux premiers moments de la distribution gaussienne  $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$ . Une telle expression est déduite d'un estimateur de variance minimale. Dans tous les autres cas (non linéaires et non gaussiens), le filtre de Kalman ne peut pas être utilisé en l'état. Des extensions de ce filtre ont été proposées pour résoudre le problème non linéaire avec des bruits toujours gaussiens tel que l'EKF et l'UKF. Cependant, dans le cas non linéaire, on ne peut pas conclure sur l'optimalité du système. On utilise alors des comparaisons avec une borne caractérisant la performance optimale : la borne de Cramer Rao *a posteriori*<sup>1</sup>.

### 3.3.2 Modèle non-linéaire avec bruit gaussien : filtre de Kalman étendu (EKF)

Lorsque le système à résoudre n'est pas linéaire, mais que les hypothèses sur les bruits sont conservées (bruits blancs additifs, gaussiens, indépendants entre eux et indépendants de la condition initiale), le système peut être approché par un modèle linéaire gaussien par linéarisation des équations d'état et de mesure qui fournissent les équations du filtre de Kalman étendu [Athans *et al.*, 1968] :

$$\begin{aligned} \mathbf{X}_k &= \mathbf{f}_k(\mathbf{X}_{k-1}) + \mathbf{w}_k \\ \mathbf{Z}_k &= \mathbf{h}_k(\mathbf{X}_k) + \mathbf{v}_k \end{aligned} \tag{3.6}$$

---

<sup>1</sup>PCRB : Posterior Cramer Rao Bound

où les fonctions  $\mathbf{f}_k$  et  $\mathbf{h}_k$  sont des fonctions non linéaires, supposées dérivables, et  $\mathbf{v}_k$  et  $\mathbf{w}_k$  des bruits blancs indépendants. Quand les fonctions  $\mathbf{f}_k$  et  $\mathbf{h}_k$  sont dérivables, on peut linéariser le système autour de l'état prédit  $\hat{\mathbf{X}}_{k|k-1}$  et de l'état courant  $\hat{\mathbf{X}}_k$ . L'extension du filtre de Kalman au cas non linéaire consiste donc à estimer les deux premiers moments et repose sur l'approximation de la loi conditionnelle par une gaussienne dont les moments sont calculés par linéarisation des fonctions  $\mathbf{f}_k$  et  $\mathbf{h}_k$ . Le plus simple est alors la linéarisation du système au premier ordre (cf. l'algorithme 2) autour de l'estimée précédente pour l'équation d'état et de la prédiction pour l'équation de mesure.

---

**Algorithme 2** Filtre de Kalman étendu

---

- **Initialisation :**  $\hat{\mathbf{X}}_{0|0} = \hat{\mathbf{X}}_0, \Sigma_{0|0} = \Sigma_0$   
pour  $k = 1, 2, \dots, K$ .
  - **Linéarisation :**

$$\mathbf{F}_k = \nabla \mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1})$$

$$\mathbf{H}_k = \nabla \mathbf{h}_k(\hat{\mathbf{X}}_{k|k-1})$$
  - **Prédiction :**

$$\hat{\mathbf{X}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{X}}_{k-1|k-1} + \mathbf{B}_k \mathbf{u}_k$$

$$\Sigma_{k|k-1} = \mathbf{F}_k \Sigma_{k-1|k-1} \mathbf{F}_k^t + \mathbf{Q}_k$$
  - **Correction :**

$$\mathbf{S}_k = \mathbf{H}_k \hat{\Sigma}_{k|k-1} \mathbf{H}_k^t + \mathbf{R}_k$$

$$\mathbf{G}_k = \hat{\Sigma}_{k|k-1} \mathbf{H}_k^t \mathbf{S}_k^{-1}$$

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{G}_k [\mathbf{Z}_k - (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1} + \mathbf{D}_k \mathbf{u}_k)]$$

$$\hat{\Sigma}_{k|k} = [\mathbf{I} - \mathbf{G}_k \mathbf{H}_k] \hat{\Sigma}_{k|k-1}$$
- 

**Commentaires**

Bien qu'intensivement utilisé en ingénierie pour traiter les problèmes non linéaires, aucun travail théorique n'a été réalisé pour justifier son utilisation dans le cas général. Néanmoins, des études valident son emploi pour la résolution des systèmes avec un bruit d'observation faible [Picard, 1991] [Oliveira, 1994]. Par contre, si les non-linéarités sont trop importantes, le filtre de Kalman étendu peut diverger. Cela est d'autant plus vrai que les équations du filtre ne prennent pas en compte l'erreur de linéarisation. Pour remédier à ce problème, quelques heuristiques peuvent être utilisées comme l'augmentation des bruits  $\mathbf{Q}_k$  et  $\mathbf{R}_k$  afin de prendre en compte cette erreur ou d'augmenter artificiellement la covariance  $\Sigma_{k|k}$  permettant de limiter la confiance dans les estimées passées. Ces méthodes nécessitent un ajustement précis et délicat des paramètres.

### 3.3.3 Modèle non-linéaire avec bruit gaussien : filtre de Kalman sans parfum (UKF)

Toujours dans le cadre de l'estimation des systèmes non linéaires, le filtre de Kalman sans parfum propose une alternative au filtre de Kalman étendu. Pour ce filtre, la densité conditionnelle est toujours supposée gaussienne et est représentée par un ensemble de points nommés  $\sigma$ -points. Ces points sont choisis de façon déterministe et sont propagés à l'aide des modèles non linéaires afin d'évaluer la moyenne  $\hat{\mathbf{X}}_{k|k}$  et la matrice de covariance  $\Sigma_{k|k}$  de l'état prédit. Le choix des points représentant la loi de filtrage, décrit



dans [Julier & Uhlmann, 1997], nécessite la connaissance des deux premiers moments de l'approximation. L'algorithme 3 décrit le filtre de Kalman sans parfum.

---

**Algorithme 3** Filtre de Kalman sans parfum
 

---

- **Initialisation :**  $\hat{\mathbf{X}}_{0|0} = \hat{\mathbf{X}}_0, \Sigma_{0|0} = \Sigma_0$   
pour  $k = 1, 2, \dots, K$ .
  - Calcul du nuage pondéré de points  $\{\mathbf{X}_{k-1|k-1}^{(n)}, \mathbf{w}_{k-1|k-1}^{(n)}\}_{n=1 \dots N}$ , ( $N$  étant le nombre de points) à partir de l'estimation précédente  $\hat{\mathbf{X}}_{k-1|k-1}, \Sigma_{k-1|k-1}$
  - **Prédiction :**

$$\hat{\mathbf{X}}_{k|k-1}^{(n)} = \mathbf{f}_k \hat{\mathbf{X}}_{k-1|k-1}^{(n)}$$

$$\hat{\mathbf{X}}_{k|k-1} = \sum_{n=1}^N \mathbf{w}_{k-1|k-1}^{(n)} \mathbf{X}_{k|k-1}^{(n)}$$

$$\Sigma_{k|k-1} = \sum_{n=1}^N \mathbf{w}_{k-1|k-1}^{(n)} [(\mathbf{X}_{k|k-1}^{(n)} - \hat{\mathbf{X}}_{k|k-1})(\mathbf{X}_{k|k-1}^{(n)} - \hat{\mathbf{X}}_{k|k-1})^t] + \mathbf{Q}_k$$

$$\hat{\mathbf{Z}}_{k|k-1}^{(n)} = \mathbf{h}_k(\hat{\mathbf{X}}_{k|k-1}^{(n)})$$

$$\hat{\mathbf{Z}}_{k|k-1} = \sum_{n=1}^N \mathbf{w}_{k-1|k-1}^{(n)} \mathbf{Z}_{k|k-1}^{(n)}$$
  - **Correction :**

$$\Sigma_{zz} = \sum_{n=1}^N \mathbf{w}_{k-1|k-1}^{(n)} [(\mathbf{Z}_{k|k-1}^{(n)} - \hat{\mathbf{Z}}_{k|k-1})(\mathbf{Z}_{k|k-1}^{(n)} - \hat{\mathbf{Z}}_{k|k-1})^t] + \mathbf{R}_k$$

$$\Sigma_{xz} = \sum_{n=1}^N \mathbf{w}_{k-1|k-1}^{(n)} [(\mathbf{X}_{k|k-1}^{(n)} - \hat{\mathbf{X}}_{k|k-1})(\mathbf{Z}_{k|k-1}^{(n)} - \hat{\mathbf{Z}}_{k|k-1})^t]$$

$$\mathbf{G}_k = \Sigma_{xz} \Sigma_{zz}^{-1}$$

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{G}_k [\mathbf{Z}_k - \hat{\mathbf{Z}}_{k|k-1}]$$

$$\Sigma_{k|k} = \Sigma_{k|k-1} \mathbf{G}_k \Sigma_{zz} \mathbf{G}_k^t$$
- 

**Commentaires**

Ce filtre conserve l'approximation de la loi de filtrage  $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$  par une loi gaussienne. Cependant, il n'y a plus de linéarisation des équations d'état et de mesure. Cela permet alors d'éviter les deux principaux inconvénients du filtre de Kalman étendu, à savoir les erreurs de linéarisation et les difficultés de calcul de l'expression analytique des matrices jacobiniennes. De plus, alors que l'EKF approxime la moyenne et la covariance de la gaussienne au premier ordre (au sens d'un développement de Taylor), nous pouvons démontrer que cette version sans parfum peut être précise jusqu'au troisième ordre. Ce filtre améliore donc les performances, du fait de la meilleure estimation des moments.

**3.3.4 Synthèse**

Lorsque le système est linéaire et gaussien, le filtre de Kalman fournit la solution optimale au problème de filtrage récursif. En revanche, dès que le modèle du système n'est plus linéaire, ce filtre n'est plus valide. Le filtre de Kalman étendu permet de pallier ce problème par une linéarisation du système par les développements de Taylor. Malheureusement, ce filtre doit se limiter au modèle présentant une faible non linéarité car sinon il peut diverger facilement. Le filtre de Kalman sans parfum permet de répondre à ce problème et donne de très bons résultats sur des applications fortement non linéaires. Malgré ses meilleures performances, le filtre de Kalman sans parfum reste limité aux hypothèses gaussiennes. Cette hypothèse est rapidement mise en défaut dans le cas général, c'est-à-dire lorsque la densité de filtrage n'est plus gaussienne (multimodale par exemple).

## 3.4 Méthodes stochastiques

Lorsque le modèle décrivant le système dynamique à estimer est non linéaire et/ou non gaussien, les méthodes non aléatoires (filtre de Kalman et extensions, méthodes numériques par maillage) échouent (divergent) pour donner une approximation précise de l'état courant des objets suivis. L'utilisation de simulations par Monte Carlo pour la construction de cette loi est une alternative.

### 3.4.1 Le filtre à particules SIR

Les filtres particuliers SIR<sup>2</sup> sont fondés sur le principe de Monte Carlo pour estimer et prédire en ligne des systèmes dynamiques. Ces filtres permettent de représenter la distribution  $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$  par un jeu d'échantillons discrets concentré en des points dénommés « particules » et pondérés par des coefficients appelés « poids des particules » (cf. l'équation 3.8).

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) \simeq \sum_{n=1}^N \pi_k^n \delta(\mathbf{X}_k - \mathbf{X}_k^n) \quad (3.7)$$

$$\pi_k^n = \frac{\pi_k^n}{\sum_{n=1}^N \pi_k^n} \quad (3.8)$$

Ces filtres consistent à faire évoluer au cours du temps ce système de particules ainsi que leur poids et reposent sur l'usage récursif du principe d'échantillonnage pondéré et une sélection par rééchantillonnage. Les particules assorties de leurs poids permettront notamment d'accéder à toutes les distributions de probabilité régissant l'équation 3.3 qui peut se réécrire comme suit :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) = C^{-1} p(\mathbf{Z}_k | \mathbf{X}_k) \sum_{n=1}^N \pi_{k-1}^n p(\mathbf{X}_k | \mathbf{X}_{k-1}^n) \delta(\mathbf{X}_{k-1} - \mathbf{X}_{k-1}^n) \quad (3.9)$$

Ce principe fut indépendamment proposé et utilisé au début des années 90 par [Moral *et al.*, 1992] [Gordon *et al.*, 1993a] [Kitagawa, 1996] [Isard & Blake, 1996] sous le nom de filtre particulière SIR.

Une illustration de la propagation des particules par l'algorithme SIR (cf. algorithme 4) est présentée en figure 3.5. Depuis, le filtre particulière SIR reposant sur les méthodes d'échantillonnage pondérées a été largement exploité dans de nombreuses applications. Citons par exemple le traitement du signal (pistage par mesure d'angles, navigation, etc...) [Gustafsson *et al.*, 2002] [Hue *et al.*, 2002] [Bréhard & Cadre, 2004], le traitement de la parole [Vermaak *et al.*, 2002], les problèmes de classification [Freitas *et al.*, 2001], le suivi d'entités en vision par ordinateur [Isard & Blake, 1998] [Okuma *et al.*, 2004], la robotique mobile [Montemerlo *et al.*, 2002] [Franck *et al.*, 2003] [Karlsson & Gustafsson, 2001].

### Commentaires

L'idée du filtre SIR consiste à mettre à jour récursivement une approximation discrète de la loi *a posteriori* qui est constituée par un nuage pondéré de particules. La mise à jour du nuage se fait par des tirages aléatoires, selon les équations induites par le modèle de

<sup>2</sup>Sequential Importance Resampling

---

**Algorithme 4** Filtre particulaire SIR (Sequential Importance Resampling)

---

- **Initialisation** : Générer  $\mathbf{X}_0^{(n)} \sim p(\mathbf{X}_0)$  avec  $\pi_0^{(n)} = \frac{1}{N}$ .  
pour  $k = 1, 2, \dots$
  - **Propagation des particules** : Générer  $\mathbf{X}_k^{(n)} \sim g(\mathbf{X}_k | \mathbf{X}_{k-1}^{(n)})$ .
  - **Mise à jour des poids** : Calculer les poids d'importance  $\pi_k^{(n)}$  pour chaque particule  $\mathbf{X}_k^{(n)}$  à partir de l'évaluation de la fonction de vraisemblance  $\pi_k^{(n)} \propto p(\mathbf{Z}_k | \mathbf{X}_k^{(n)})$ .
  - **Normaliser les poids d'importance** :  $\pi_k^{(n)} = \frac{\pi_k^{(n)}}{\sum_{n=1}^N \pi_k^{(n)}}$
  - **Rééchantillonnage** : Tirer  $N$  particules parmi  $\mathbf{X}_k^{(n)}$  avec la probabilité  $\pi_k^{(n)}$ .
- Le nuage de particules obtenu à l'instant  $k$  approxime la distribution *a posteriori*  
 $p(\mathbf{X}_k | \mathbf{Z}_{1:k}) \simeq \{\mathbf{X}_k^{(n)}, \pi_k^{(n)}\}_{n=1}^N$
- 

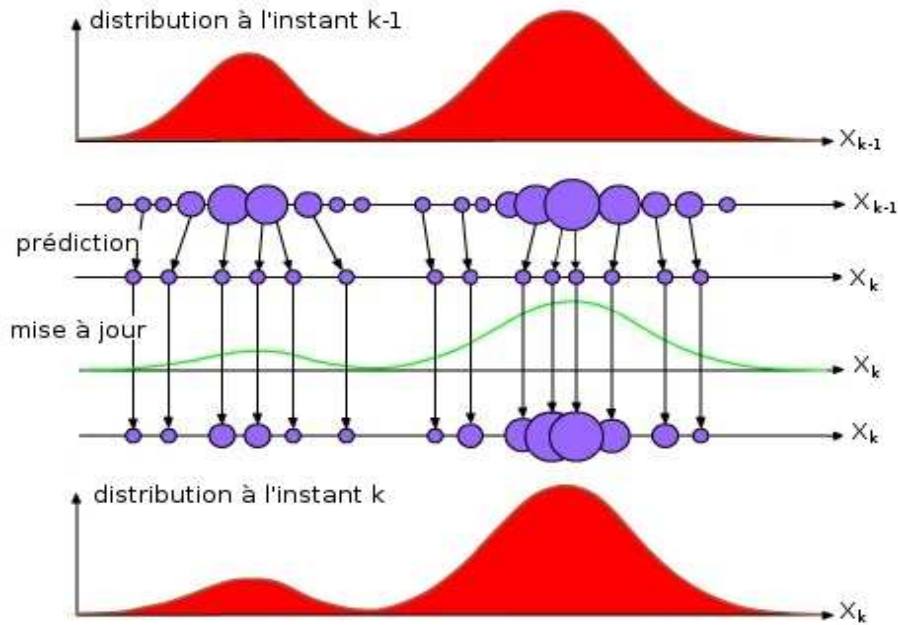


FIG. 3.5 – Propagation des particules par l'algorithme SIR - figure tirée de [Smith, 2007].

vraisemblance considéré. La principale difficulté de mise en oeuvre de ce filtre, est d'avoir une très bonne connaissance du système analysé afin de pouvoir générer le maximum de particules utiles lors de l'évaluation de l'état caché du système. En effet, l'efficacité de ce filtre est liée à la capacité de la loi d'évolution à prédire des particules dans les régions de forte vraisemblance.

### 3.4.2 L'échantillonneur à particules MCMC

L'échantillonneur à particules MCMC <sup>3</sup> est une méthode de rééchantillonnage qui ne propage plus les particules de manière parallèle comme les filtres de la famille SIR, mais construit le jeu de particules par une chaîne de Markov (chaîne d'hypothèses), liées l'une

---

<sup>3</sup>Monte Carlo Markov Chain

à l'autre par un processus markovien : la  $k+1$ -ème hypothèse est générée à partir de la  $k$ -ème hypothèse.

L'échantillonneur MCMC a été proposé ces dernières années afin d'offrir une approche alternative efficace au problème de suivi multi-objets à partir d'une image vidéo. L'idée principale est de remplacer l'étape traditionnelle d'échantillonnage pondérée par une étape d'échantillonnage basée sur les méthodes de Monte Carlo par chaînes de Markov. L'échantillonneur MCMC, explore l'espace d'état par une chaîne de Markov à l'ordre 1, selon le traditionnel algorithme de Metropolis-Hasting [Hastings, 1970], où la transition d'un échantillon  $\mathbf{X}$  au suivant  $\mathbf{X}^*$  est assurée par une fonction de proposition  $q(\mathbf{X}^*|\mathbf{X})$ .

### L'algorithme de Metropolis-Hasting

Introduit dans sa forme la plus simple par Metropolis en 1953 puis généralisé par Hastings en 1970, l'algorithme portant aujourd'hui leur nom est une procédure permettant de simuler un échantillon d'une distribution, univariée ou multidimensionnelle, de variable aléatoire.

L'algorithme 5 s'appuie sur une distribution de proposition  $q(\mathbf{X}^*|\mathbf{X})$  qui permet de générer facilement de nouvelles propositions  $\mathbf{X}^*$  à partir de  $\mathbf{X}$  sous la forme d'une chaîne de Markov ergodique. Le calcul du taux d'acceptation  $\alpha$  permet d'accepter ou de rejeter la proposition  $\mathbf{X}^*$ . Notons qu'il est nécessaire d'éliminer les premiers échantillons (*burn-in*) de la chaîne de Markov ayant tendance à biaiser la distribution de probabilité (cf. la figure 3.6).

---

#### Algorithme 5 L'algorithme de Metropolis-Hasting pour un échantillonneur MCMC

---

##### Initialisation :

Initialiser la chaîne de Markov  $\mathbf{X}_k^{(0)}$  avec l'état d'un échantillon tiré au hasard  $\{\mathbf{X}_{k-1}^{(n)}\}_{n=1}^N$  à l'instant  $k - 1$  et son poids  $\pi(\mathbf{X}_k^{(0)}) \propto p(\mathbf{Z}_k|\mathbf{X}_k^{(0)})$ .

##### Metropolis-Hasting :

**pour**  $n = 1$  à  $N + N_B$

Tirer une proposition  $\mathbf{X}_k^*$  de la distribution  $q(\mathbf{X}_k^*|\mathbf{X}_k)$  avec  $\mathbf{X}_k = \mathbf{X}_{k-1}^{(n)}$

Evaluer sa vraisemblance :  $\pi(\mathbf{X}_k^*) \propto p(\mathbf{Z}_k|\mathbf{X}_k^*)$

Calculer le taux d'acceptation :

$$\alpha = \min(1, \frac{\pi(\mathbf{X}_k^*)}{\pi(\mathbf{X}_k)})$$

Ajouter un  $n$ ème élément de la chaîne  $\mathbf{X}_k^{(n)} = \mathbf{X}^*$  avec la probabilité  $\alpha$ ,  
sinon la chaîne conserve l'élément précédent  $\mathbf{X}_k^{(n)} = \mathbf{X}_k$ .

**fin du pour**

**Suppression :** Supprimer les  $N_B$  premiers échantillons de la chaîne.

Le nuage de particules obtenu à l'instant  $k$  approxime la distribution *a posteriori*  $p(\mathbf{X}_k|\mathbf{Z}_k)$ .

---

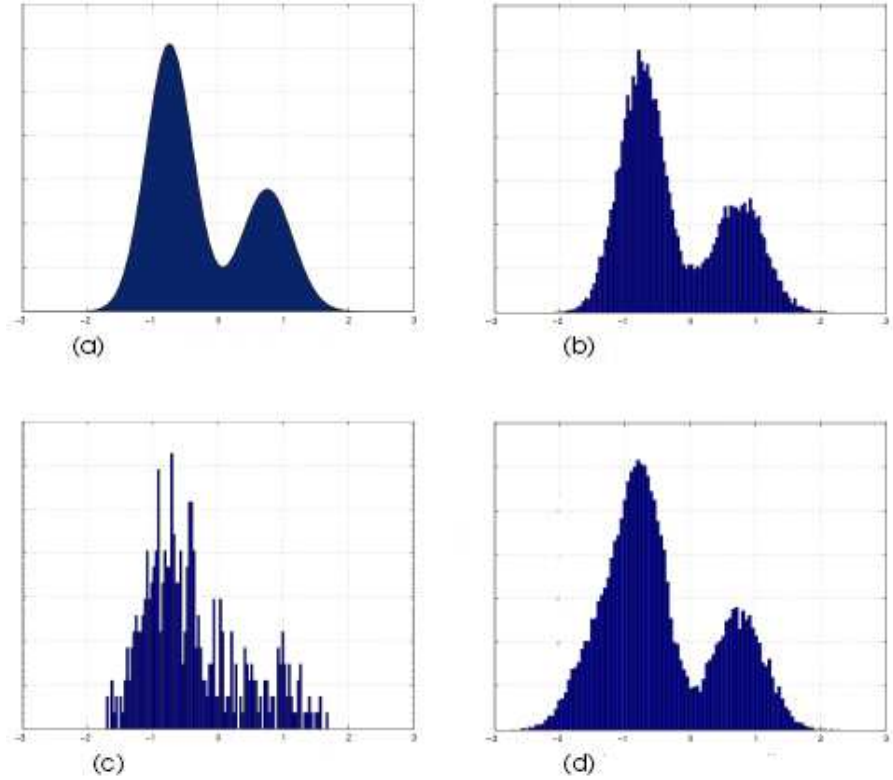


FIG. 3.6 – Approximation d’une distribution cible  $p(X)$  par un échantillonneur MCMC. En (a), la distribution cible  $p(X)$  est présentée. En (b), le résultat de l’approximation MCMC par Metropolis-Hasting de la distribution cible pour un nombre suffisamment grand d’échantillons  $N$  et après élimination des  $N_B$  premiers échantillons (*burn-in*). La distribution est bien approximée. En (c), résultat de l’approximation MCMC par Metropolis-Hasting de la distribution cible mais avec un nombre insuffisant d’échantillons  $N$ . La distribution est alors mal approximée. Enfin, en (d), résultat de l’approximation MCMC par Metropolis-Hasting de la distribution cible pour un nombre suffisamment grand d’échantillons  $N$  mais sans élimination des  $N_B$  premiers échantillons. La distribution se retrouve alors biaisée par le point de départ (l’initialisation) de notre chaîne - figure extraite de [Smith, 2007].

### Commentaires

En pratique, il est important de noter qu’un échantillonneur MCMC n’est pas un filtre car il construit la nouvelle distribution *a posteriori*  $p(\mathbf{X}_k|\mathbf{Z}_k)$  à partir d’une unique hypothèse tirée aléatoirement dans la distribution  $p(\mathbf{X}_{k-1}|\mathbf{Z}_{k-1})$ .

### 3.4.3 Le filtre à particules MCMC

Egalement appelés Markov Chain Monte Carlo Particle Filter (MCMC PF), ces filtres furent introduits par [Khan *et al.*, 2004] puis par [Smith, 2007] pour le suivi d’un grand nombre d’objets. Le principe est de remplacer l’échantillonnage par importance par un échantillonnage de Metropolis. L’algorithme 6 résume le fonctionnement de ce filtre intégrant un échantillonneur de Metropolis-Hasting où la construction de la loi  $p(\mathbf{X}_k|\mathbf{Z}_k)$  est

obtenue à partir de l'ensemble des particules tiré de la loi  $p(\mathbf{X}_{k-1}|\mathbf{Z}_{k-1})$ .

---

#### Algorithme 6 L'algorithme du filtre MCMC

---

##### Initialisation :

Initialiser la chaîne de Markov  $\mathbf{X}_k^{(0)}$  avec l'état d'un échantillon tiré au hasard  $\{\mathbf{X}_{k-1}^{(n)}\}_{n=1}^N$  à l'instant  $k-1$  et son poids  $\pi(\mathbf{X}_k^{(0)}) \propto p(\mathbf{Z}_k|\mathbf{X}_k^{(0)})$ .

##### Metropolis-Hasting :

**pour**  $n = 1$  à  $N + N_B$

Tirer une proposition  $\mathbf{X}_k^*$  de la distribution  $p(\mathbf{X}_k|\mathbf{Z}_{k-1})$  avec  $\mathbf{X}_k = \mathbf{X}_{k-1}^{(n)}$

Evaluer sa vraisemblance :  $\pi(\mathbf{X}_k^*) \propto p(\mathbf{Z}_k|\mathbf{X}_k^*)$

Calculer le taux d'acceptation :

$$\alpha = \min(1, \frac{\pi(\mathbf{X}_k^*)}{\pi(\mathbf{X}_k)})$$

Ajouter un énième élément de la chaîne  $\mathbf{X}_k^{(n)} = \mathbf{X}_k^*$  avec la probabilité  $\alpha$ ,  
sinon la chaîne conserve l'élément précédent  $\mathbf{X}_k^{(n)} = \mathbf{X}_k$ .

##### fin du pour

**Suppression :** Supprimer les  $N_B$  premiers échantillons de la chaîne.

Le nuage de particules obtenu à l'instant  $k$  approxime la distribution *a posteriori*  $p(\mathbf{X}_k|\mathbf{Z}_k)$ .

---

#### Commentaires

L'idée du filtre MCMC consiste à explorer l'espace d'état de manière itérative en créant une chaîne de Markov. En effet, le nuage de particules n'est plus généré parallèlement comme les filtres SIR, mais séquentiellement, de manière chaînée, afin de fabriquer à un instant  $k$  le nouveau nuage de particules. La principale difficulté de mise en oeuvre de ce filtre, concerne la convergence vers l'équilibre de la chaîne de Markov, qui peut être lente, surtout quand la loi cible est difficile à approcher et que l'initialisation de la chaîne est « malheureuse ». De plus, et afin d'éviter de biaiser l'approximation de la fonction de densité de probabilité, le choix du nombre de particules  $N_B$  à éliminer est à déterminer.

### 3.4.4 Synthèse

Depuis le début des années 90, l'utilisation des méthodes séquentielles de Monte Carlo pour l'analyse de systèmes non linéaires et/ou non gaussiens ont fait l'objet de recherches de plus en plus actives tant en statistique théorique que dans les domaines applicatifs. Ces méthodes apportent une nouvelle solution au problème de filtrage non linéaire dont les performances dépassent celles des algorithmes paramétriques dans le cas général. Deux méthodes différentes ont été présentées dans cette section. La première repose sur l'utilisation de l'échantillonnage pondéré portée par la famille des filtres particuliers de type

SIR. La seconde repose sur l'utilisation des chaînes de Markov portée par la famille des filtres particuliers de type MCMC. Ces deux méthodes séquentielles de Monte Carlo consistent à chercher la loi *a posteriori*  $p(\mathbf{X}_k|\mathbf{Z}_{1:k})$  par une exploration de l'espace d'état parallèle ou séquentielle à l'aide de  $N$  hypothèses appelées particules. Le principal avantage d'une exploration d'état MCMC par rapport à une propagation SIR est la possibilité d'effectuer des mouvements marginalisés selon un sous-espace de dimension  $d$ . Par contre la propagation des particules SIR est mieux adapté au calcul en parallèle, donc à des machines multi-processeurs.

### 3.5 Méthode d'association pour le suivi multi-objets

Dans les deux sections précédentes, nous nous sommes intéressés à l'analyse du problème de filtrage. Deux grandes familles d'algorithmes (paramétriques et stochastiques) ont été citées avec pour chacune plusieurs algorithmes décrits. Le choix de la méthode de filtrage dépend principalement des cibles à suivre, donc du modèle d'évolution permettant de prédire leur mouvement. Ces algorithmes cherchent à déterminer l'état caché des cibles situées dans le champs de vue du ou des capteur(s) à partir des nouvelles mesures disponibles à chaque nouvelle itération  $k$ .

En pratique, plusieurs mesures sont souvent disponibles à chaque itération ; chacune de ces mesures pouvant être issue de l'entité suivie ou avoir été générée par le processus d'élaboration des mesures sans provenir des objets (également appelé fausses alarmes). La prise en compte des fausses alarmes dans la mise à jour de la loi de filtrage perturbe la qualité du résultat et peut amener à une perte de l'entité suivie. Il est donc primordial d'être capable de distinguer, à chaque instant  $k$ , quelles sont les mesures informatives sur l'état du système à estimer de celles qui sont des fausses alarmes. C'est le problème de l'association de données (en anglais : *data association*) (cf. la figure 3.7).

Le problème de l'association de données est encore plus important dans un cadre de filtrage multi-objets. En effet, il doit permettre de déterminer si une mesure est une fausse alarme ou dans le cas contraire, être capable de déterminer à quel objet elle appartient. De nombreux travaux ont été menés sur cette problématique dont les plus connus sont ceux proposés dans l'ouvrage de BarShalom [Bar-Shalom & Li, 1995].

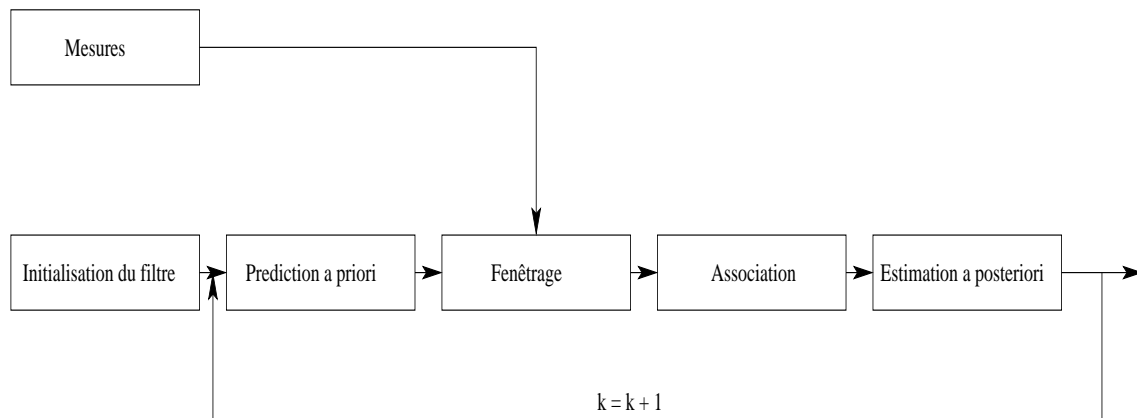


FIG. 3.7 – Le module de suivi.



### 3.5.1 Fenêtre de validation des mesures

Pour limiter la complexité des calculs et éviter des associations aberrantes, une première étape consiste à réaliser une approximation appelée « fenêtre de validation des mesures » (cf. la figure 3.7), qui consiste à délimiter, à partir de la mesure prédite  $\tilde{\mathbf{X}}_{(k|k-1)}$ , un volume de l'espace d'observation, à l'extérieur duquel la cible a une probabilité très faible de se trouver. Les mesures extérieures au volume sont éliminées car elles sont statistiquement trop éloignées de la position prédite. Seules les mesures les plus probables sont conservées, elles sont dites validées.

Dans le cadre d'un suivi multi-objets mono-capteur, deux hypothèses d'association peuvent être considérées. Elles définissent les règles d'association entre les mesures validées et les pistes.

**H1** : Une mesure peut provenir d'un objet ou d'une fausse alarme.

**H2** : A un instant donné, un objet peut être à l'origine d'aucune ou d'au plus une mesure.

La combinaison des hypothèses **H1** et **H2** traduit le fait que les associations possibles décrites dans la matrice d'association  $\mathbf{A}$  (avec  $i \in \{1, \dots, O\}$  objets et  $j \in \{1, \dots, m_k\}$  mesures) sont exhaustives et mutuellement exclusives.

#### Modèle gaussien

Pour calculer le test de validation  $T(z)$ , il faut supposer que la mesure  $\mathbf{Z}_k$  provient de la cible et vérifie :

$$p(\mathbf{Z}_k | \tilde{\mathbf{X}}_{k-1}) = \mathcal{N}(\mathbf{Z}_k, \tilde{\mathbf{X}}_{k|k-1}; \mathbf{S}_k) \quad (3.10)$$

où  $\mathcal{N}(\mathbf{Z}, \mu; \mathbf{S})$  désigne la densité de probabilité gaussienne de variable  $\mathbf{Z}$ , de moyenne  $\mu$  et de covariance  $\mathbf{S}$ . L'état prédit  $\tilde{\mathbf{X}}_{k|k-1}$  et la covariance prédite  $\mathbf{S}_k$  sont disponibles à l'instant  $k$ .

Le test de validation s'obtient en imposant un seuil minimal à la densité *a priori* de la mesure  $p(\mathbf{Z}_k | \tilde{\mathbf{X}}_{k-1})$ , d'où le résultat :

$$\begin{cases} T(z) = 1 & \text{si } \tilde{\mathbf{Z}}_k^t \mathbf{S}_k^{-1} \tilde{\mathbf{Z}}_k \leq \gamma \\ T(z) = 0 & \text{si } \tilde{\mathbf{Z}}_k^t \mathbf{S}_k^{-1} \tilde{\mathbf{Z}}_k > \gamma \end{cases} \quad (3.11)$$

avec

$$\tilde{\mathbf{Z}}_k = (\mathbf{Z}_k - \tilde{\mathbf{X}}_{k|k-1}) \quad (3.12)$$

et

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^t + \mathbf{R}_k \quad (3.13)$$

Le seuil  $\gamma$  est fixé en choisissant la taille de la fenêtre de validation de telle sorte que la probabilité de trouver la mesure correcte à l'intérieur de la fenêtre de validation soit égale à  $P_g$ .  $\gamma$  est lié à  $P_g$  par la relation :

$$P_g = p(\chi_{nz}^2 \leq \gamma) \quad (3.14)$$

où  $\chi_{nz}^2$  désigne une variable qui suit une loi du  $\chi^2$  à  $nz$  degré de liberté.

#### Modèle non gaussien

Dans le cas du filtrage particulaire, la densité  $p(\tilde{\mathbf{X}}_k | \mathbf{Z}_{k-1})$  n'est pas connue analytiquement. Cependant, le nuage de particules pondérées  $\{\mathbf{x}_{k-1}^{(n)}, \mathbf{w}_{k-1}^{(n)}\}_{n=1 \dots N}$  permet d'approcher la densité de vraisemblance de la nouvelle observation. En effet, en connaissant



les échantillons  $\{\tilde{\mathbf{x}}_{k-1}^{(n)}\}_{n=1\dots N}$  obtenus par prédiction du nuage  $p(\mathbf{X}_{k-1}|\mathbf{Z}_{1:k-1})$  on peut poser :

$$p(\tilde{\mathbf{X}}_k|\mathbf{Z}_{k-1}) \simeq \sum_{n=1}^N \tilde{w}_{k-1}^{(n)} p(\mathbf{Z}_k|\tilde{\mathbf{x}}_{k-1}^{(n)}) \quad (3.15)$$

A partir de cette expression, il est possible de se ramener à une hypothèse gaussienne en calculant analytiquement les deux premiers moments de l'estimation  $p(\tilde{\mathbf{X}}_k|\mathbf{Z}_{k-1})$ , et donc de construire une fenêtre ellipsoïdale de validation des mesures.

Cependant, afin de conserver le caractère non gaussien de la distribution  $p(\tilde{\mathbf{X}}_k|\mathbf{Z}_{k-1})$ , une autre solution consiste à construire numériquement ces régions de validation en utilisant par exemple les méthodes non paramétriques à noyaux (KDE).

### 3.5.2 Méthodes d'association paramétriques

#### Filtre à association de données probabilistes (PDAF)

L'augmentation importante de la puissance de calcul a rendu populaire les méthodes utilisant toutes les observations voisines d'une piste pour sa mise à jour.

Le PDAF [Bar-Shalom & Fortmann, 1988] est une extension du filtre de Kalman dédiée au filtrage d'un objet unique évoluant dans un environnement bruité. Cet algorithme se décompose donc en deux étapes, prédiction et correction. La prise en compte des différentes hypothèses d'association est faite au travers de calcul de l'innovation (cf. la figure 3.8).

Cette formulation revient à approcher une somme de densités de gaussiennes par une unique densité gaussienne.

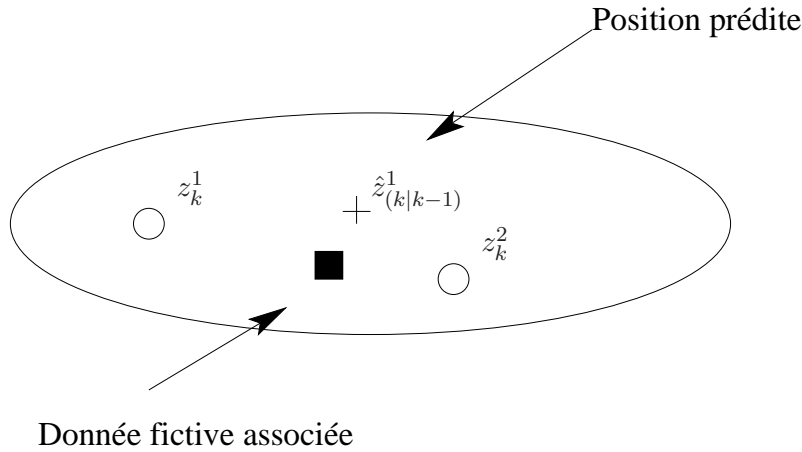


FIG. 3.8 – L'Association par PDAF consiste à construire un estimateur qui prend en compte toutes les observations de la fenêtre, donc toute l'information contenue dans celle-ci. Sur cette figure, on peut voir que la piste est associée avec une donnée « fictive » calculée en pondérant chaque donnée présente dans la fenêtre de validation.

#### Filtre joint à association de données probabilistes (JPDAF)

Le JPDAF [Bar-Shalom & Fortmann, 1988] [Gorji *et al.*, 2007] est une extension du PDAF pour la poursuite multi-objets. Le JPDAF suppose normalement un nombre connu

$O$  d'objets suivis. Mais, il est facile de l'adapter au suivi d'un nombre inconnu d'objets en ajoutant un module de gestion de piste afin de gérer la création et la destruction de piste.

Le JPDAF permet de renforcer la robustesse du PDAF dans les zones critiques où deux fenêtres se chevauchent (cf. la figure 3.9) en apportant l'information multi-cibles. Le PDAF est alors modifié de la façon suivante :

- ▷ si les interactions des fenêtres sont vides, on remet à jour de simples filtres de poursuite PDAF, indépendamment les uns des autres ;
- ▷ si les interactions ne sont pas vides, on examine toutes les hypothèses envisageables sans dissocier les différentes pistes. Cela revient à établir une matrice d'association  $A$  permettant de décrire l'hypothèse qu'une mesure  $z_k^j$  reçue à l'instant  $k$  provienne de la piste  $x_k^i$ .

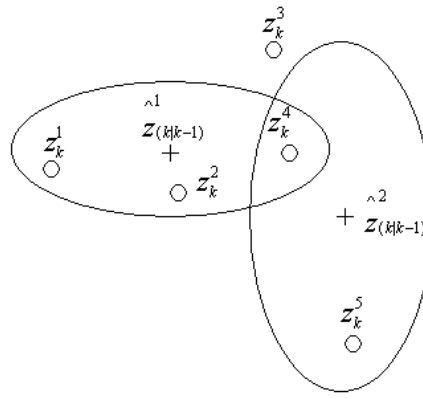


FIG. 3.9 – Chevauchement de deux fenêtres de validation pour deux objets et cinq mesures.

### 3.5.3 Méthodes d'association non paramétriques

#### Méthode du plus proche voisin (PPV)

Cette méthode [Bar-Shalom & Fortmann, 1988] est la plus simple et la plus utilisée pour répondre au problème de l'association de données. C'est une méthode purement séquentielle qui permet de trouver l'association la plus probable entre une observation et une piste existante ou entre une nouvelle observation et une nouvelle piste. L'idée principale est de trouver la plus petite distance qui sépare une piste  $x_k^i$ , parmi toutes les observations  $z_k^j$  qui appartiennent à la région de validité de  $x_k^i$  (cf. la figure 3.10).

Le principal inconvénient de cette méthode du plus proche voisin tient au fait que la décision sur l'association est immédiate et irrévocable : l'hypothèse la plus probable est considérée comme vraie. Son principal avantage est son faible coût calculatoire. Ce type de méthode est donc particulièrement adapté aux environnements contenant à chaque nouvel instant  $k$ , peu d'observations et surtout peu de fausses alarmes. Le suivi d'un nombre inconnu d'objets s'effectue, comme pour le JPDAF, en ajoutant un module de gestion de piste. Une analyse des performances de cette méthode est présentée dans [Bar-Shalom & Fortmann, 1988].

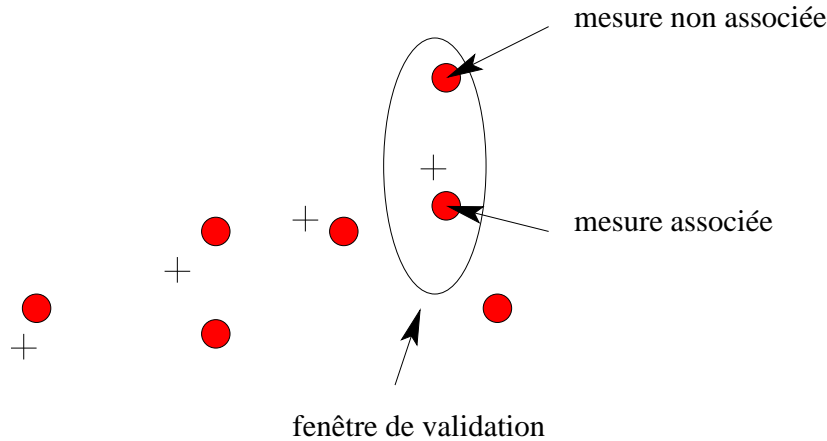


FIG. 3.10 – Exemple d'une association par PPV.

### Filtre à hypothèses multiples (MHT)

Dans cette méthode [Read, 1979], le nombre d'objets est inconnu mais supposé supérieur à un, et pouvant croître ou décroître au cours du temps. Le principe du MHT est de considérer systématiquement toutes les hypothèses d'association possibles et de créer une piste pour chacune de ces hypothèses. L'idée est d'attendre que de nouvelles observations capteurs lèvent l'ambiguïté et permettent de décider quelles hypothèses sont à abandonner et les quelles sont à garder (cf. la figure 3.11).

Bien que le coût de mise en oeuvre de cet algorithme soit élevé, les performances du filtre MHT sont significatives. En effet, ce dernier gère naturellement la création, destruction et la fusion des pistes. En revanche, le nombre d'hypothèses suivies croît exponentiellement à chaque itération avec les nouvelles observations capteurs. Il est donc nécessaire d'éliminer les hypothèses dont les probabilités sont trop faibles. Ce nombre d'hypothèses devra donc être assez élevé afin de diminuer la complexité de l'algorithme mais sans exagérer afin de ne pas éliminer trop rapidement une hypothèse de probabilité à un instant donné qui pourrait se révéler intéressante par la suite.

### 3.5.4 Méthodes stochastiques

Face à un modèle non linéaire, il est possible d'utiliser des méthodes d'association de données basées sur les méthodes de Monte Carlo. Ces méthodes permettent de capter les non-linéarités du système à travers un nuage d'hypothèses (ou particules) sans nécessiter d'approximation gaussienne par exemple. En effet, approchant la loi  $p(\mathbf{X}_{0:k}|\mathbf{Z}_{1:k})$ , le nuage de particules transporte également les informations sur les associations passées.

En définitif, les principales difficultés qui sont liées à l'utilisation des méthodes séquentielles de Monte Carlo résident dans la définition de comment :

- ▷ calculer la fenêtre de validation permettant de valider les mesures.
- ▷ modéliser la vraisemblance  $p(\mathbf{Z}_k|\mathbf{X}_k)$ .

Ces dernières années, plusieurs solutions d'association de données basées sur ces méthodes ont été proposées. On peut notamment citer :

- ▷ Monte Carlo Joint Probabilistic Data Association (MCJPDA) [Karlsson & Gustafsson, 2001] [Vermaak *et al.*, 2005],

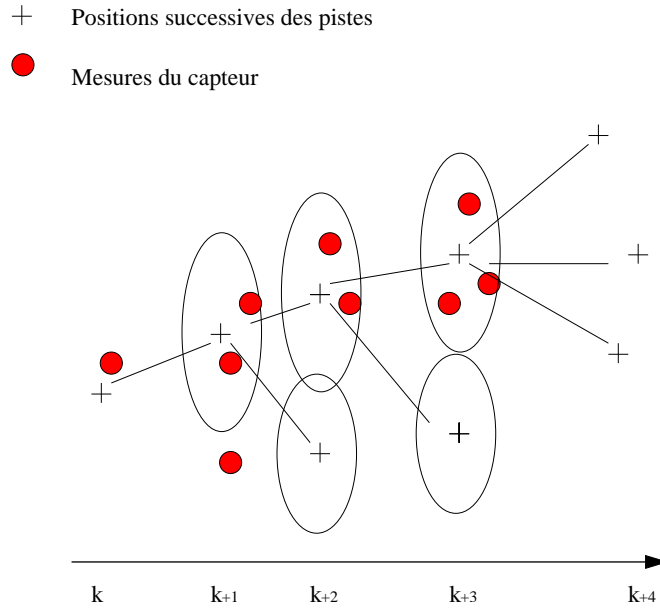


FIG. 3.11 – Exemple d'associations par MHT.

- ▷ Rao-Blackwellized Monte Carlo Data Association (RBMCD) [Sarkka *et al.*, 2004],
- ▷ Markov Chain Monte Carlo Data Association (MCMCD) [Oh *et al.*, 2004].

#### 3.5.5 Synthèse

Dans cette section, nous nous sommes intéressés au problème d'association de données dans le cadre d'un suivi mono-capteur multi-objets. Ce problème apparaît lorsque plusieurs mesures et/ou plusieurs pistes sont disponibles à chaque itération. Une première étape consiste à valider ou non les mesures en fonction de la fenêtre de validation. Dans le cas où une unique mesure est conservée, les algorithmes de filtrage présentés dans les deux premières sections de ce chapitre peuvent être appliqués directement. Dans le cas contraire, l'utilisation des algorithmes présentée dans cette section (PPV, JPDAF, MHT, RBMCD, MCMCD...) permet de modéliser l'incertitude sur l'origine des mesures et de résoudre le problème d'association de données.

**Notre algorithme d'association des données et de fenêtrage repose sur une stratégie non paramétrique, fondée sur la méthode du plus proche voisin (PPV) car notre algorithme de détection fournit peu d'ambiguïté sur les observations comme le ferait un capteur radar par exemple, qui donne plusieurs mesures d'un seul objet.**

## 3.6 Extension du problème de l'association de données à la fusion multi-capteurs

### 3.6.1 Introduction

La fusion de données multi-capteurs correspond à la volonté d'utiliser simultanément plusieurs sources de données, ayant des caractéristiques éventuellement différentes, afin

d'obtenir une nouvelle information de meilleure qualité.

Dans une première partie, cette section présente la notion de fusion de données en rappelant les principales approches utilisées dans le cadre de la perception d'environnement incertains.

Dans un second temps, l'extension du suivi bayésien au cas multisensoriel nous amène à introduire la notion d'architecture de fusion qui est primordial lors du développement d'une application de poursuite multi-objets multi-capteurs. Les derniers travaux concernant la fusion entre un capteur laser et une caméra sont énumérés et classés suivant l'architecture choisie.

Enfin, un tableau comparatif des performances des différentes architectures est proposé à la fin de cette section.

#### 3.6.2 La fusion de données : Buts et approches classiques

Le problème de la combinaison d'informations à partir de plusieurs sources se rencontre dans de nombreux champs d'application souvent liés au besoin de vouloir percevoir un environnement à partir de capteurs plus ou moins fiables et précis. En se basant sur les principaux travaux effectués ces dernières années en robotique mobile, nous rappelons dans ce paragraphe un panel non exhaustif des principales techniques de fusion rencontrées. Ces techniques sont ici appliquées à la combinaison et à l'intégration de données hétérogènes provenant de plusieurs sources. C'est pourquoi nous les appelons « approches classiques ».

Si l'on se restreint au domaine de l'intelligence artificielle et de la robotique, on peut identifier principalement trois grandes classes d'approches de fusion :

- ▷ celle basée sur la théorie de l'évidence [Denoeux & Govaert, 1996] [Lefevre *et al.*, 2004].
- ▷ celle basée sur la théorie des ensembles flous [Jouannin, 1999].
- ▷ celle basée sur les modèles probabilistes [Douillard *et al.*, 2007] [Premibida *et al.*, 2007].

Une comparaison de ces trois stratégies de fusion est proposée et commentée dans [Chateau, 1999].

#### Commentaires

Dans le domaine du contrôle de processus dynamique et surtout dans le domaine de la poursuite de cibles en mouvement, les filtres bayésiens se sont révélés performants, comme par exemple dans la poursuite de cibles aériennes autant dans le domaine militaire que civil. Ces filtres ont également trouvé leurs applications dans de nombreux domaines, comme la navigation robotique et la fusion multi-capteurs.

**Pour ces raisons, c'est cette méthode qui a été choisie dans cette thèse pour résoudre le problème de la fusion de données laser/vidéo.**

#### 3.6.3 Architectures multi-capteurs classiques dans le cadre d'un filtre bayésien

Le choix architectural du système multi-capteurs permet de définir à quel niveau seront effectués le filtrage et la maintenance des pistes par rapport à l'étape de fusion des

données. La fusion multi-capteurs étant un domaine large d'application (suivi d'entité mobile, navigation, localisation...), nous nous sommes restreints dans le cadre de nos travaux, à recenser exclusivement tout article s'attachant à détecter et/ou suivre une personne en position dans un contexte urbain à l'aide d'un système de capteur combinant laser et vidéo. Dans le cadre de la détection, puis du suivi de piéton à l'aide d'un capteur laser et vidéo, nous avons recensé quatre principales architectures dans huit articles ou revues depuis 2005.

#### Architecture de fusion série

Le principe de la fusion série est très simple. Dans un premier temps, le capteur laser permet de détecter des objets (classifiés ou pas comme piéton) par un algorithme de détection. Dans un second temps, tous ces objets sont projetés dans le repère caméra en tant que simples hypothèses et deviennent des zones d'intérêts de recherche en vision par un algorithme de classification. La vision valide ou non ces hypothèses en les identifiant en tant que piétons ou non. Cette méthode est illustrée par la figure 3.12.

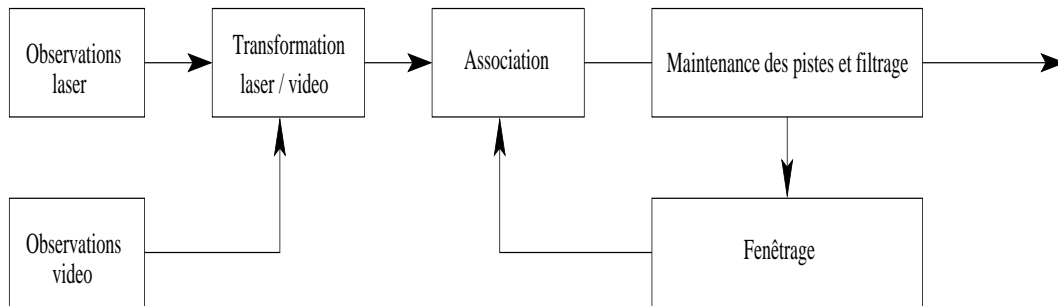


FIG. 3.12 – Architecture de fusion série entre un capteur laser et un capteur vidéo. Les observations laser alimentent la recherche vidéo avant la phase de poursuite multi-pistes.

Szarvas et al. [Szarvas *et al.*, 2006] exploitent de cette manière la complémentarités des deux capteurs. En effet, le capteur laser permet d'aider à localiser les objets susceptibles d'être intéressants dans l'image vidéo. Ces objets sont ensuite classifiés en utilisant un algorithme de vision basé sur des réseaux de neurones convolutifs (Convolutional Neural Networks).

Cui et al. [Cui *et al.*, 2005] commencent par détecter les jambes des piétons dans l'image laser avant de projeter leurs positions dans l'image vidéo. A partir d'un gabarit de hauteur *a priori* d'un piéton ( $\sim 1.70m$ ), l'image vidéo confirme ou pas la présence d'un piéton à partir de l'analyse d'un histogramme de couleur.

#### Architecture de fusion centralisée

La deuxième méthode consiste à considérer les observations des différents capteurs de manière globale, pour réaliser l'association de données, la maintenance et le filtrage des pistes. La fusion d'estimation centralisée a pour but d'estimer le vecteur d'état  $X_k$ , en utilisant l'ensemble des mesures disponibles dans une centrale de calculs unique. Pour ce type de fusion, on utilise l'ensemble des observations fournies par les  $M$  capteurs utilisés avant de prendre une décision finale. Les différentes mesures peuvent être synchronisées ou non, pondérées ou non suivant leur provenance (capteur laser, vidéo, etc.). Cette méthode est illustrée pour un capteur laser et vidéo par la figure 3.13.

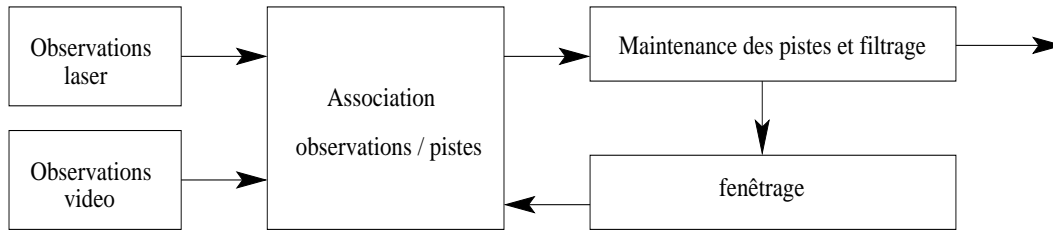


FIG. 3.13 – Architecture de fusion centralisée entre un capteur laser et un capteur vidéo. Les observations des deux capteurs sont traitées dans le module d'association "globale" avant la phase de poursuite multi-pistes.

Spinello et al. [Spinello *et al.*, 2009] proposent de détecter et de suivre des piétons et des véhicules en combinant des informations issues de deux algorithmes indépendants basés sur l'utilisation d'un capteur laser et d'une caméra vidéo. L'algorithme de détection laser classe les données laser en utilisant un segmenteur fondé sur les Champs Conditionnels Aléatoires (ou Conditional Random Fields CRF). L'algorithme de détection vidéo classe les objets en utilisant une extension du modèle implicite de forme (Implicit Shape Model ISM). La combinaison de ces deux sources d'informations est réalisée par un filtre de Kalman étendu (Extended Kalman Filter EKF).

Thornton et al. [Thornton *et al.*, 2008] proposent une approche de détection et de suivi de piéton utilisant des informations issues d'un capteur laser 3D et d'une caméra vidéo infrarouge. L'algorithme de détection laser permet de détecter les différents objets en comptant sur leur information de hauteur. En passant par un gabarit *a priori* d'une personne, l'algorithme classe ou pas un objet en tant que piéton. L'algorithme de vision compte sur la différence thermique entre une personne et le reste de la scène afin de détecter les piétons. La combinaison de ces deux sources d'informations est réalisée par une approche multi-niveaux qui combine au mieux les avantages et les inconvénients des deux capteurs.

### Architecture de fusion décentralisée

La troisième méthode consiste à faire une poursuite multi-cibles au niveau des différents capteurs. Chaque capteur possède donc sa propre logique d'association de données, de maintenance et de filtrage des pistes. La fusion décentralisée (ou distribuée) permet donc d'estimer le vecteur d'état  $\mathbf{X}_k$  en utilisant l'ensemble des états estimés en provenance de plusieurs unités de calculs. La fusion combine alors les états estimés  $(\hat{\mathbf{X}}_{1,k}, \dots, \hat{\mathbf{X}}_{M,k})$  des  $M$  capteurs dans une nouvelle estimation  $\hat{\mathbf{X}}_k$  du vecteur d'état. Chaque estimation  $\hat{\mathbf{X}}_{m,k}$  est observée par une mesure  $\mathbf{Z}_{m,k}$ , donnée par le capteur  $m$ . Chaque capteur traite respectivement son observation avant de transmettre la valeur estimée au centre commun de fusion. Cette méthode est illustrée pour un capteur laser et vidéo par la figure 3.14.

Kurazume et al. [Kurazume *et al.*, 2008] proposent un algorithme de suivi basé sur des données issues d'un réseau de caméras et de capteurs lasers. Sa méthode repose sur une fusion décentralisée des informations basée sur un nouveau filtre particulier combinant les méthodes SIR (Sequential Importance Resampling) et MCMC (Markov Chain Monte Carlo).

Goya et al. [Goya *et al.*, 2008] proposent un algorithme de détection et de suivi de véhicule à partir d'un capteur laser et d'une caméra vidéo formalisé par un filtre particu-



### 3.6. EXTENSION DU PROBLÈME DE L'ASSOCIATION DE DONNÉES À LA FUSION MULTI-CAPTEURS

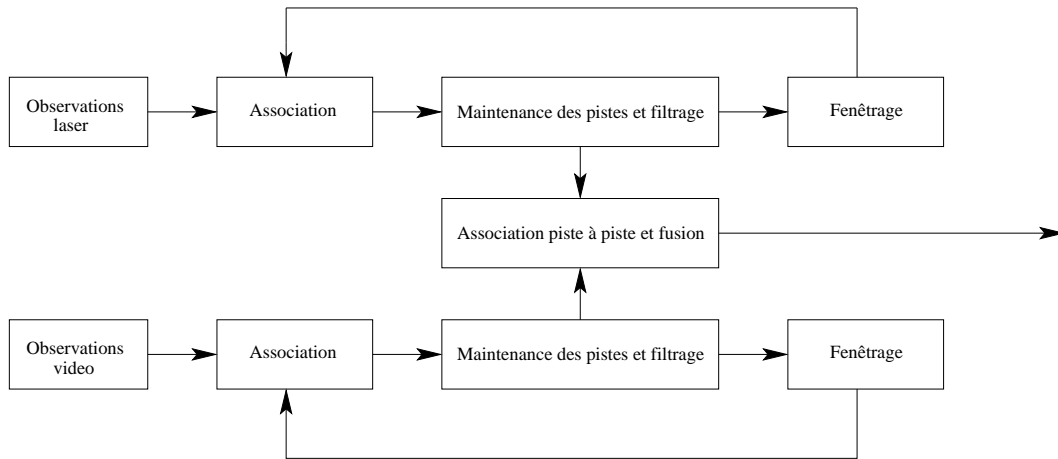


FIG. 3.14 – Architecture de fusion décentralisée entre un capteur laser et un capteur vidéo. Chaque capteur fait sa propre poursuite multi-objets, les pistes des deux capteurs sont ensuite fusionnées.

laire. Le principe de sa méthode repose sur la manière dont il combine puis échantillonne les poids d'importance des différentes particules (appelé échantillonnage multi-sources) provenant des deux capteurs.

#### Architecture de fusion hybride

La quatrième méthode consiste à combiner la deuxième et la troisième méthode. On peut noter une prépondérance de la fusion hybride sur une combinaison des données fournies par une fusion série avec une fusion centralisée ou décentralisée. Cette méthode est illustrée par la figure 3.15.

Monteiro et al. [Monteiro *et al.*, 2006] proposent un algorithme de détection et de suivi de piéton basé sur la détection d'objets dans l'espace laser avant de les classifier comme piéton ou non par un algorithme de vision. Il propose aussi de classifier indépendamment les piétons dans l'espace laser par un classifieur GMM<sup>4</sup> et dans l'espace vidéo par un classifieur AdaBoost. La combinaison de ces informations est effectuée par une somme bayésienne.

Ngako et al. [Ngako-Pangop *et al.*, 2008] présentent une approche bayésienne permettant de fusionner des informations provenant d'un capteur laser et d'une caméra vidéo. Sa principale contribution est d'ajouter la modélisation des différentes observations par un modèle autoregressif lors de la classification bayésienne des piétons.

#### 3.6.4 Synthèse

Le choix de l'architecture est fonction de l'application et des données qui sont disponibles. Dans le cas du développement d'une application de poursuite multi-objets multi-capteurs sur un véhicule automobile par exemple, le temps de traitement de l'algorithme ainsi que le taux de transfert des données entre les différents capteurs seront un critère de choix déterminant. Les architectures « fusion série » et « fusion centralisée » demandent moins de temps de calcul mais sont plus sensibles aux moindres défauts des capteurs

<sup>4</sup>Gaussian Mixture Model



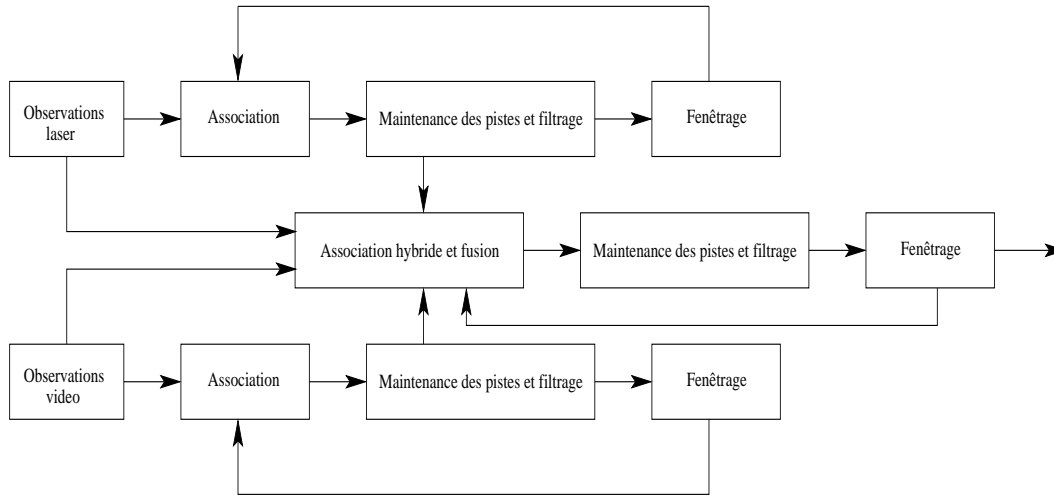


FIG. 3.15 – Architecture de fusion hybride entre un capteur laser et un capteur vidéo. Chaque capteur fait sa propre poursuite multi-objets. Les observations et les pistes des deux capteurs sont traitées dans le module d'association "globale" avant la phase de poursuite multi-pistes.

Architecture	Avantages	Inconvénients
Fusion série	Robuste aux fausses alarmes ; utilisation directe des algorithmes d'association de données mono-capteur.	Très sensible aux pannes capteurs et aux qualités des détections laser.
Fusion centralisée	Utilisation directe des algorithmes d'association de données mono-capteur ; temps de calcul.	Moins robuste aux pannes capteurs.
Fusion décentralisée	Robuste aux pannes capteurs ; meilleure estimation de l'incertitude de position.	Besoin de calculer la fusion des deux covariances ; association moins précise.
Fusion hybride	Robuste aux pannes capteurs ; meilleure estimation de la position et de son incertitude	Complexité plus importante de l'algorithme ; temps de calcul plus élevé.

TAB. 3.1 – Principaux avantages et inconvénients des quatres architectures de fusion entre un laser et une caméra.

tandis que les architectures « fusion décentralisée » et « fusion hybride » sont plus gourmandes en temps de calcul mais moins sensibles aux défauts des capteurs et plus précises dans l'estimation de l'incertitude en position. Enfin, l'adaptation des algorithmes « classiques » d'association de données est plus directe dans le cas des architectures « fusion série » où cette association est faite de manière séquentielle, c'est-à-dire à chaque nouvelle arrivée de mesures.

## 3.7 Conclusion

L'estimation de l'état caché d'un système dynamique à partir d'observations bruitées est un problème compliqué qui fait l'objet de recherches de plus en plus actives depuis le début des années 90. Lorsque les données sont reçues séquentiellement, l'utilisation des algorithmes itératifs est nécessaire.

Le filtre bayésien est un algorithme itératif permettant de déterminer l'estimation de la loi *a posteriori* de l'état caché du système étudié. Dans un cadre linéaire et gaussien, le filtre de Kalman permet d'obtenir l'estimation optimale de cette loi. Dans un cadre « légèrement » non linéaire et gaussien, les différentes extensions de Kalman permettent d'obtenir une très bonne estimation de la loi recherchée. Dans un cadre général non-linéaire et non-gaussien, les méthodes séquentielles de Monte Carlo permettent d'apporter une solution à l'estimation de la loi *a posteriori*. Leur principe consiste à approcher la loi *a posteriori* par une somme pondérée de mesures de Dirac appelées particules. La mise à jour de ce nuage de particules se fait en deux étapes : une étape d'échantillonnage pondéré faisant évoluer le nuage, suivie d'une étape de rééchantillonnage permettant de ne conserver que les particules les plus pertinentes.

Outre les problèmes de non-linéarité du système, le filtrage peut être perturbé par une mauvaise association de données due à la présence de fausses alarmes (observations non issues des entités suivies) ou d'observations manquantes (défaillance de l'algorithme de détection ou des capteurs, occultation des entités suivies...). Les algorithmes d'association de données sont utilisés pour lever l'ambiguïté sur l'origine des mesures. Dans un cadre gaussien, une extension du filtre de Kalman, le JPDAF, est la méthode la plus couramment utilisée. Dans le cas non-linéaire, les méthodes séquentielles de Monte Carlo sont plus adaptées.

Face à un système donné, le choix de l'algorithme de filtrage et de l'association de données dépend des caractéristiques du modèle suivi (linéaire ou non, gaussien ou non, environnement de mesures bruitées ou non). L'étude de la modélisation du système est donc comme toujours très importante et sera proposée dans le chapitre 4.

Enfin, nous nous sommes intéressés à la fin de ce chapitre au problème de fusion de données dans le cadre d'un suivi multi-objets multi-capteurs. Après un rapide survol des différentes théories de fusion de données, nous avons présentés dans le cadre d'une fusion de données laser/vidéo bayésienne, les avantages et inconvénients de quatre architectures. Rappelons que dans un cadre bayésien, les techniques d'associations de données telles que PPV, JPDAF, MHT, RBMCDA, MCMCDA... restent valables, et sont utilisées pour résoudre le problème de fusion de données.



---

## Chapitre 4

# Algorithmes proposés pour la détection et le suivi de piéton

Ce chapitre décrit dans un premier temps un nouvel algorithme pour le problème de détection des piétons à partir de mesures délivrées par un télémètre laser. A la lumière de l'état de l'art présenté au chapitre 2, nous proposons une méthode d'extraction d'« objets-piétons » spécifique à la géométrie humaine décrite dans une image de profondeur comme celle délivrée par un capteur IBEO. Cette méthode qui est non paramétrique nous permet de calculer la fonction de vraisemblance d'une image laser à partir des caractéristiques géométriques *a priori* (taille et disposition spatiale des amas de points) d'un piéton.

Dans le cas d'un capteur multi-plans, une combinaison décentralisée des nappes est proposée afin d'améliorer le taux de détection et de diminuer le nombre de fausses alarmes. Cette combinaison repose sur une méthode similaire à celle proposée pour l'extraction d'« objets-piétons » à la différence que la discrimination effectuée par la fonction noyau ne se focalise plus sur la géométrie d'un piéton mais sur la dispersion spatiale des « objets-piétons » détectés dans l'étape précédente.

Dans un second temps, un algorithme de suivi de piéton à partir de mesures délivrées par un ou plusieurs capteurs est étudié. Dans le cadre d'un suivi multi-piétons, nous nous plaçons dans un contexte le plus général possible, où aucune hypothèse gaussienne sur les bruits d'état et de linéarité sur leur mouvement n'est effectuée. A la lecture du chapitre 3, notre choix s'est porté sur l'utilisation d'un filtre particulier SIR qui est une méthode suffisamment générique pour être adaptée au suivi d'entités telles que les piétons. Lors d'un suivi multi-cibles, il est nécessaire d'être capable de déterminer à quels objets appartiennent les mesures. C'est le problème de l'association des données. Ce problème est exactement le même lorsque nous parlons de fusion des données. En effet, résoudre un tel système équivaut également à déterminer à quels objets appartiennent les mesures provenant des différents capteurs. Nous proposons donc une méthode d'association générique, fondée sur les méthodes non paramétriques à noyau permettant de formaliser la combinaison des mesures provenant d'un ou plusieurs capteurs.

## 4.1 Détection puis suivi de piéton à partir d'un capteur multi-plans laser

Dans cette section, nous décrivons l'algorithme de détection puis de suivi de piéton à partir d'une image laser. Cet algorithme est composé de différentes briques (cf. la figure 4.1) qui doivent permettre à partir d'une image laser de fournir les positions et les vitesses relatives de tous les piétons situés dans le champ de vue du capteur.

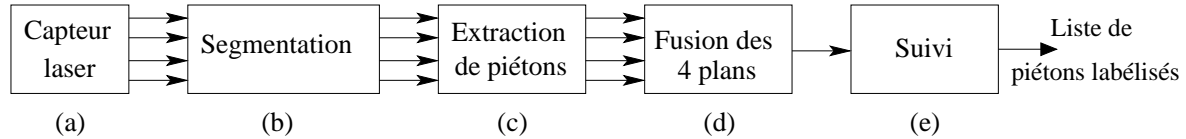


FIG. 4.1 – Algorithme de détection puis de suivi de piéton à partir d'un capteur multi-plans laser. La figure 4.2 illustre ces différentes étapes.

Chaque étape de cet algorithme est détaillée de la manière suivante :

- ▷ la brique de segmentation permettant de regrouper les points laser dans chaque plan sous forme de segments et d'amas de points est présentée dans la section 4.1.1,
- ▷ la brique permettant d'isoler les piétons du reste des objets détectés dans l'étape précédente est détaillée dans la section 4.1.2,
- ▷ la brique permettant d'exploiter les quatre plans du capteur avant de prendre une décision finale sur la classification ou non des objets comme des piétons est exposée dans la section 4.1.3,
- ▷ la brique de suivi permettant de suivre et d'identifier dans le temps les piétons en position et vitesse avec leurs différentes incertitudes associées est présentée dans la section 4.1.4.

En dehors de la brique de segmentation qui repose sur l'algorithme de Arras *et al.* [Arras, 2003], toutes les autres briques présentées dans ce chapitre ont été développées durant cette thèse.

### 4.1.1 Segmentation

L'extraction d'observations à partir de données capteur est la première étape fondamentale de tout algorithme de suivi d'objets. Le critère principal de segmentation d'une image de profondeur est l'évaluation de la distance entre deux points consécutifs  $P_i$  et  $P_{i+1}$ . L'objectif de cette étape est de regrouper, à chaque nouvelle acquisition du capteur et pour chaque obstacle de la scène observée, les points appartenant à un même obstacle en tenant compte de l'imprécision du capteur utilisé. Pour extraire des segments ou des droites, l'algorithme que nous avons choisi pour notre application fait partie des principaux algorithmes présentés et évalués dans [Nguyen *et al.*, 2005]. Cet algorithme appelé « line regression algorithm » développé par Arras *et al.* [Arras, 2003] propose de construire les différents clusters sous deux formes géométriques : des segments et des amas de points (*beacons*).

Cet algorithme se décompose en trois étapes principales :

- ▷ regroupement en cluster de tous les points 2D qui respectent un certain critère de distance [Vestli, 1995],

#### 4.1. DÉTECTION PUIS SUIVI DE PIÉTON À PARTIR D'UN CAPTEUR MULTI-PLANS LASER

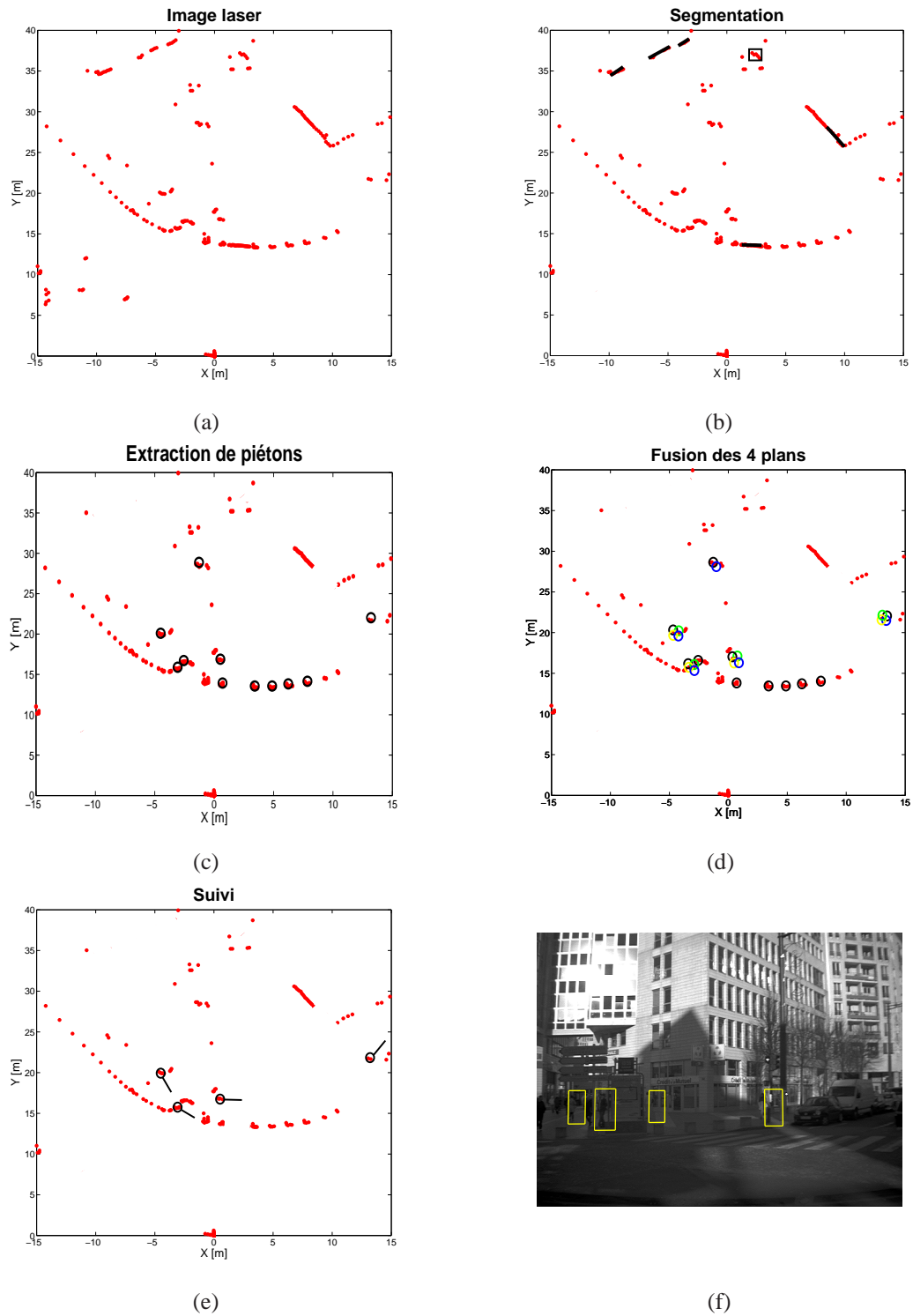


FIG. 4.2 – Illustration des différentes étapes de l'algorithme de détection puis de suivi de piéton à partir d'un capteur multi-plans laser. En bas, les piétons détectés qui sont reprojétés dans l'image vidéo.

- ▷ calcul par régression linéaire si les clusters peuvent être modélisés sous forme d'une ou plusieurs lignes,
- ▷ garder sous forme d'amas de points les clusters constitués d'un nombre minimum de points voisins en connexité et en distance.

### Extraction de droite

Le problème standard de la régression linéaire est de déterminer les paramètres de la droite permettant d'approximer au mieux un nuage de points en minimisant l'erreur commise seulement dans la dimension de la variable dépendante comme par exemple suivant  $\rho$ . L'algorithme de segmentation développé par Arras permet de déterminer les paramètres de la droite (cf. la figure 4.3) qui passe par tous les points les plus pertinents en minimisant l'erreur dans les deux dimensions, c'est-à-dire dans le cas de coordonnées polaires  $\rho$  et  $\theta$ . La régression linéaire est donc de meilleure qualité car elle estime la droite en fonction de l'erreur commise dans les deux dimensions simultanément. L'équation d'une droite polaire de paramètre  $r$  et  $\alpha$  s'écrit

$$\rho \cos(\theta - \alpha) - r = 0 \quad (4.1)$$

Les valeurs  $r$  et  $\alpha$  tirées de (4.1) minimisant l'erreur au sens des moindres carrés peuvent s'exprimer

$$\alpha = \frac{1}{2} \arctan \left\{ \frac{-2 \sum (\bar{y} - \bar{y}_i)(\bar{x} - \bar{x}_i)}{\sum [(\bar{y} - \bar{y}_i)^2(\bar{x} - \bar{x}_i)^2]} \right\} \quad (4.2)$$

$$r = \bar{x} \cos \alpha + \bar{y} \sin \alpha \quad (4.3)$$

avec les moyennes

$$\bar{x} = \frac{1}{n} \sum x_i \quad \bar{y} = \frac{1}{n} \sum y_i \quad (4.4)$$

Les variances et covariances des paramètres  $r$  et  $\alpha$  s'expriment alors

$$\sigma_\alpha^2 = \frac{1}{(D^2 + N^2)^2} \sum [N(\bar{x} \cos \theta_i - \bar{y} \sin \theta_i - \rho_i \cos 2\theta_i) - D(\bar{x} \sin \theta_i + \bar{y} \cos \theta_i - \rho_i \cos 2\theta_i)]^2 \sigma_{\rho_i}^2 \quad (4.5)$$

$$\sigma_r^2 = \sum \left[ \frac{1}{n} \cos(\theta_i - \alpha) + \frac{\partial \alpha}{\partial \rho_i} (\bar{y} \cos \alpha - \bar{x} \sin \alpha) \right]^2 \sigma_{\rho_i}^2 \quad (4.6)$$

$$\sigma_{\alpha r}^2 = \sum \frac{\partial \alpha}{\partial \rho_i} \frac{\partial r}{\partial \rho_i} \sigma_{\rho_i}^2 \quad (4.7)$$

avec

$$N = \frac{2}{n} \sum \sum P_i P_j \cos Q_i \sin Q_j - \sum P_i^2 \sin Q_i \quad (4.8)$$

$$D = \frac{1}{n} \sum \sum P_i P_j \cos(Q_i + Q_j) - \sum P_i^2 \cos 2Q_i \quad (4.9)$$



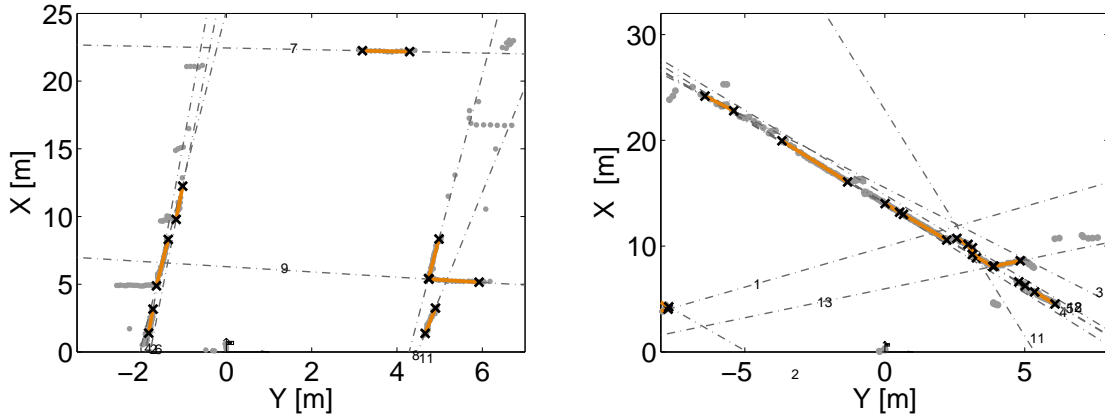


FIG. 4.3 – Résultats d'une extraction de segments dans une image laser.

### Extraction d'amas de points (ou *beacons*)

Après l'étape d'extraction de droite, le traitement se poursuit sur les clusters non approximés par des segments en les agrégeant en amas de points (cf. la figure 4.4) voisins en connexité et en distance. Chaque amas de points est décrit en coordonnées polaires avec sa matrice de covariance associée

$$x_b = [\rho_b \ \theta_b]^t \quad (4.10)$$

$$C_b = \begin{bmatrix} (\sigma_\rho^2)^{-1} & 0 \\ 0 & (\sigma_\theta^2)^{-1} \end{bmatrix} \quad (4.11)$$

Pour chaque *beacon*, les équations qui permettent d'obtenir la coordonnée de profondeur  $\rho_b$  et sa variance  $\sigma_\rho^2$  sont

$$\rho_b = \sigma_\rho^2 \sum_i \frac{\rho_i}{\sigma_{\rho_i}^2} \quad (4.12)$$

$$\sigma_\rho^2 = \sum_i \sigma_{\rho_i}^{-2} \quad (4.13)$$

avec  $\rho_i$  et  $\sigma_{\rho_i}$  qui sont respectivement la profondeur et la variance associées d'un point laser 2D d'indice  $i$ . Pour chaque *beacon*, les équations qui permettent d'obtenir la coordonnée de l'angle  $\theta_b$  et sa variance  $\sigma_\theta^2$  sont

$$\theta_b = \frac{1}{n_b} \sum_i \theta_{b_i} \quad (4.14)$$

$$\sigma_\theta^2 = \left( \frac{1}{4} \Delta_\theta (n_b + 1) \right)^2 \quad (4.15)$$

avec  $n_b$  le nombre de points minimum de l'amas,  $\Delta_\theta$  l'angle de résolution du capteur laser et  $\theta_{b_i}$  l'angle d'un point laser 2D d'indice  $i$ .

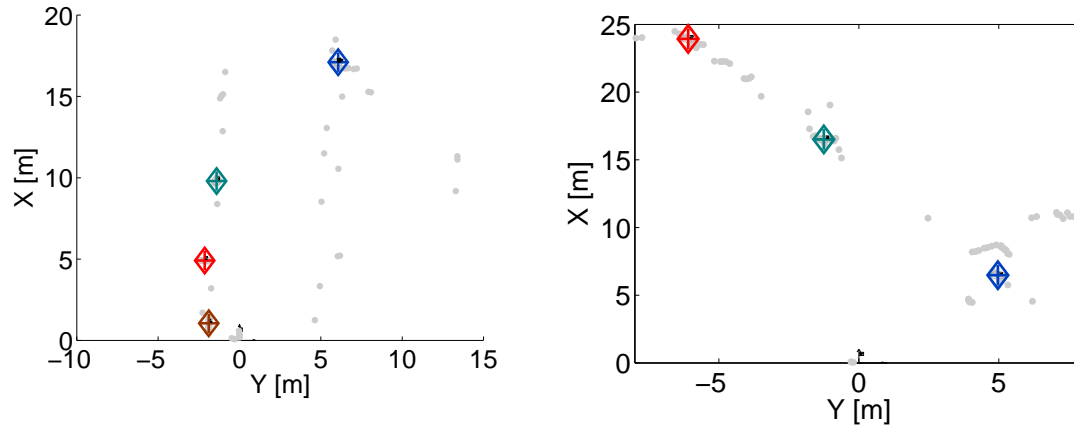


FIG. 4.4 – Résultats d'une extraction d'amas de points dans une image laser.

### Filtrage des segments et *beacons*

Sans connaissance *a priori* du nombre d'obstacles dans la scène observée, la segmentation de l'image laser permet de filtrer tous les segments et *beacons* présentant une taille géométrique importante et qui peuvent être relatifs à des murs, véhicules ou panneaux afin de ne garder que les « petits » objets pouvant être assimilés à un piéton. En conséquence, tous les impacts agrégés à un de ces grands segments ou *beacons* sont définitivement éliminés de l'image laser.

Ce filtrage permet donc de simplifier l'image de profondeur en diminuant le nombre de points qui la constitue afin de simplifier le travail de l'algorithme d'extraction d'« objets-piétons ».

#### 4.1.2 Extraction de piétons

Si pour un être humain il est aisé de reconnaître des piétons, « personnes qui circulent à pied », du reste des objets situés dans une scène, une telle tâche n'est pas aisée à assimiler pour un système de perception artificielle tel que celui examiné ici.

Nous proposons de traiter dans cette section, le problème de la détection des piétons à l'aide d'un capteur laser. Son principal avantage est la possibilité d'obtenir l'information de profondeur naturellement et de manière précise (cf. le chapitre 2), mais en contre-partie les données fournies contiennent peu d'informations sur les objets observés, en particulier pour les objets de petite taille tels que les piétons par exemple. La figure 4.5 illustre très bien ce problème avec un scan laser d'une scène où sont situés plusieurs piétons. A la question, « Où sont les piétons ? », la réponse risque d'être difficile à donner, même par un être humain. C'est pourtant ce challenge que nous allons essayer de relever en tenant compte des petites caractéristiques géométriques laissées par un piéton dans une image laser.

En effet, dans une image laser, un piéton est assimilé à un amas de points dont la répartition spatiale fluctue en fonction de trois variabilités :

- ▷ sa hauteur, qui influe sur le nombre de nappes qui le coupe,
- ▷ sa distance le séparant du capteur, qui influe sur le nombre de faisceaux qui le coupe,
- ▷ sa posture, c'est-à-dire s'il est de face (marchant en direction du capteur) ou de côté (traversant devant le capteur), qui influe sur sa largeur d'exposition au faisceau

laser.

Ces trois caractéristiques influent directement sur la disposition et le nombre d'impacts laser qu'un « objet-piéton » produit dans une image de profondeur.

La méthode présentée permet de détecter la position des amas de points dont la dimension est approximativement connue. Cette méthode est formalisée dans un cadre probabiliste où les détections sont effectuées à partir du calcul de la fonction de vraisemblance (densité de probabilité) de l'image laser.

### Estimation de densité

Les différentes approches rencontrées dans la littérature peuvent se classer en deux catégories :

- ▷ les méthodes paramétriques,
- ▷ les méthodes non paramétriques.

Les méthodes paramétriques supposent que la distribution des données suit une forme fonctionnelle spécifique régie par un petit nombre de paramètres comme par exemple, une loi gaussienne.

Les méthodes non paramétriques n'établissent aucune supposition quant à la distribution des données. Elles construisent des modèles locaux centrés autour de chacune des données.

Dans le cas de la détection d'« objets-piétons », il s'agit d'objets non rigides possédant une variation de taille et de forme importante et qui sont situés dans un environnement urbain encombré (nombreux obstacles). Nous choisissons donc d'utiliser une méthode non paramétrique.

### L'approche non paramétrique

Les méthodes non paramétriques permettent d'estimer la probabilité d'un point en fonction de la répartition spatiale des autres points présents dans son voisinage immédiat. Pour déterminer le voisinage immédiat on peut :

- ▷ soit fixer le nombre de points  $K$  habitant cette région, c'est la méthode des  $K$  plus proche voisins (KNN).
- ▷ soit fixer l'étendue  $V$  de la région, c'est la méthode des noyaux (KDE aussi appelée « fenêtres de Parzen »).

Nous proposons d'utiliser une méthode de détection non paramétrique fondée sur les fonctions à noyau qui permettent une sélection élégante des points laser les plus pertinents en construisant une carte d'extremum (cf. la figure 4.6) traduisant le niveau de confiance donné à un « objet-piéton ». L'analyse de cette carte par l'estimateur du maximum de vraisemblance permet la classification de chaque point laser comme le centre de gravité d'un piéton ou non. Notons que la méthode développée est non-supervisée car nous considérons n'avoir aucune connaissance préalable sur le nombre d'objets qui sont situés dans l'environnement observé par le capteur laser.

### Calcul de la fonction de vraisemblance d'une image de points

Soit  $I$  l'image de profondeur étudiée. On note  $\mathbf{Z} \doteq \{\mathbf{z}_k\}_{k=1,\dots,N_s}$  le vecteur composé de  $N_s$  points laser 2D qui n'ont pas été filtrés par le module de segmentation dans un plan laser. Nous définissons une variable de Bernoulli  $w_k \in \{w_1, w_2\}$  pouvant prendre

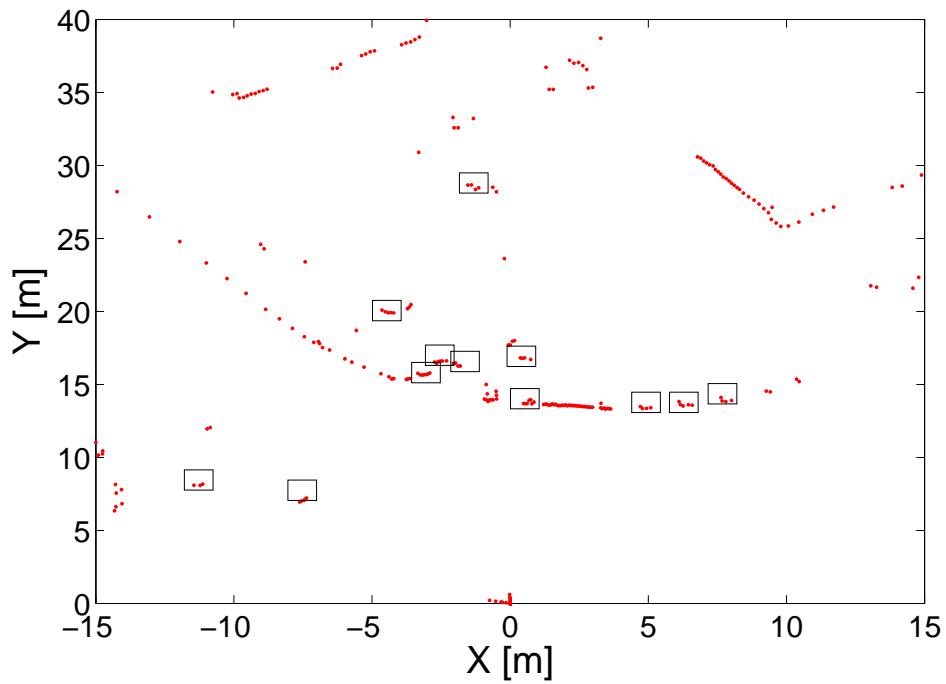


FIG. 4.5 – Exemple d’une image laser comportant plusieurs piétons à détecter. Dans une image de profondeur un « objet-piéton » est représenté par un amas de points répartis géométriquement en fonction de sa largeur, de sa distance avec le capteur et de sa posture. Sans vérité terrain, un homme risque de détecter tous les carrés bleus.

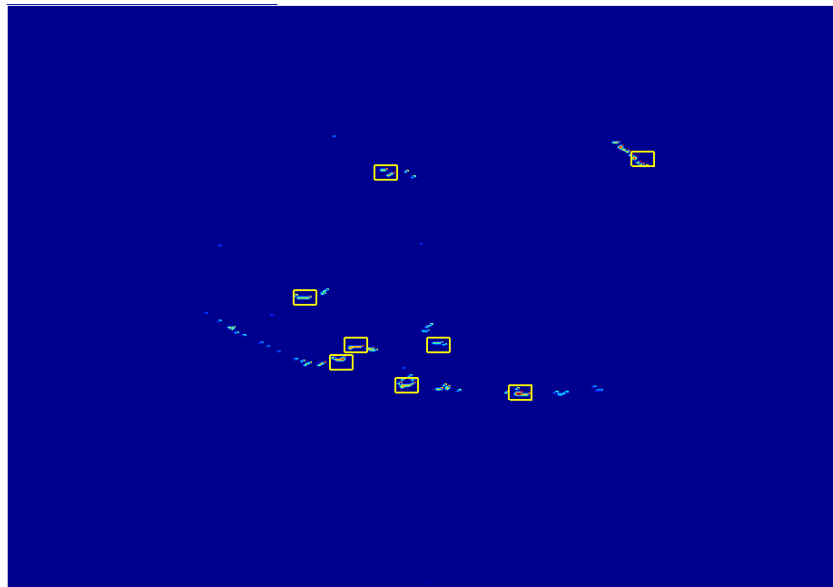


FIG. 4.6 – Exemple d’une extraction d’« objets-piétons » à partir de l’image laser ci-dessus, obtenue à partir d’une densité non paramétrique à noyaux comportant les caractéristiques géométriques d’un piéton (cf. l’équation 4.19).

deux états :  $w_k = w_1$  si l'événement associé est classifié comme le centre de gravité d'un piéton et  $w_k = w_2$  dans tous les autres cas. La fonction de vraisemblance  $p(\mathbf{Z}|w_k)$  permet de calculer la probabilité qu'un point laser appartient au centre de gravité d'un piéton. Cette fonction de vraisemblance étant par définition inconnue et imprévisible à l'avance, une fonction non paramétrique utilisant une estimation fondée sur les fonctions à noyaux (ou encore méthode de Parzen-Rozenblatt) est utilisée pour calculer la vraisemblance :

$$p(\mathbf{Z}|w_k) = \frac{1}{Nbpts(k)} \sum_{i=1}^{N_s} \varphi(\mathbf{z}_k, \mathbf{z}_i) \quad (4.16)$$

où  $N_s$  représente le nombre total de points 2D présents dans une nappe laser.  $Nbpts$  représente le nombre théorique d'impacts décrivant un piéton dans l'image de profondeur.

Pour chaque point, on calcule une distance entre sa position  $\mathbf{z}_k$  et la position de chacun des autres points  $\mathbf{z}_i$  présents dans l'image. La somme de toutes les distances ainsi calculées forme alors un poids associé à la position du point 2D courant. La fonction noyau permettant le calcul de cette somme  $\varphi(\mathbf{z}_k, \mathbf{z}_i)$  est alors définie par :

$$\varphi(\mathbf{z}_k, \mathbf{z}_i) = \exp\left[-\frac{1}{2} \cdot d_c(\mathbf{z}_k, \mathbf{z}_i)\right] \quad (4.17)$$

La distance  $d_c$  utilisée est une distance de Mahalanobis définie par :

$$d_c(\mathbf{z}_k, \mathbf{z}_i) = (\mathbf{z}_k - \mathbf{z}_i) \Sigma_\varphi^{-1} (\mathbf{z}_k - \mathbf{z}_i)^T \quad (4.18)$$

Avec  $\Sigma_\varphi$ , la matrice de covariance associée aux deux composantes géométriques liées à un piéton (largeur et épaisseur) dans une image de profondeur :

$$\Sigma_\varphi = \begin{bmatrix} \sigma_{largeur}^2 & 0 \\ 0 & \sigma_{épaisseur}^2 \end{bmatrix} \quad (4.19)$$

Le choix de  $\sigma_{largeur}$  et  $\sigma_{épaisseur}$  permet d'ajuster la taille d'un amas de points que l'on veut détecter. Enfin, la fonction qui permet de pondérer la probabilité qu'un amas de points donné par la vraisemblance  $p(\mathbf{Z}|w_k)$  soit bien un piéton est définie par :

$$Nbpts(k) = \frac{L_g}{D_k \cdot \tan \theta_k} \quad (4.20)$$

Cette expression prend en compte les dimensions d'un piéton, sa largeur  $L_g$ , la résolution angulaire  $\theta$  du capteur en fonction de la distance  $D_k$  séparant l'obstacle du véhicule.

### Localisation d'un amas de points (ou « objet-piéton »)

Une fois la vraisemblance de l'image laser définie, il convient de rechercher la position de l'amas de points possédant la plus forte probabilité. La position dominante  $\hat{\mathbf{z}}_m$  possédant le poids maximum associé à l'amas, notée  $\hat{\mathbf{Z}}_k|k$  est donnée par l'estimateur du maximum de vraisemblance comme position d'un « objet-piéton » dans l'image.

$$\hat{\mathbf{z}}_m = \hat{\mathbf{Z}}_k|k = \arg \max_k (p(\mathbf{Z}|w_k = w_1)) \quad (4.21)$$

### Localisation de plusieurs amas de points (ou d'« objets-piétons »)

Dans le cas d'une image dans laquelle plusieurs zones sont présentes la relation (4.21) fournit la position du centre de gravité de l'amas le plus représentatif. En considérant que la taille  $\kappa$  de cet amas est grossièrement connue, une liste de points  $L^m$  (en considérant l'indice  $m$  associé au numéro de l'amas courant détecté) est éliminée.

$$\mathbf{L}^m = \{N_s \in I | d_c(\hat{\mathbf{z}}_m, \mathbf{z}_i) < \kappa\} \quad (4.22)$$

Le seuil  $\kappa$  est fixé en choisissant la taille de la fenêtre de validation de telle sorte que la probabilité de trouver les mesures laser correctes soit égale à  $P_\kappa$ .  $\kappa$  est lié à  $P_\kappa$  par la relation :

$$P_\kappa = p(\chi_{nz}^2 \leq \kappa) \quad (4.23)$$

où  $\chi_{nz}^2$  désigne une variable qui suit une loi du  $\chi^2$  à  $nz$  degré de liberté.

Une nouvelle liste de points candidats  $\mathbf{L}^{m+1}$  est alors constituée. Les relations (4.16) et (4.21) sont réitérées selon l'algorithme 7 tant que l'estimateur du maximum de vraisemblance contient une valeur supérieure à un seuil de confiance  $\delta=[0,1]$ .

---

**Algorithme 7** Algorithme de détection de piétons avec un estimateur non paramétrique à noyau

---

**Entrées :** Liste des points composés des mesures laser 2D, qui n'ont pas été filtrés par le module de segmentation :  $\mathbf{Z} = \{\mathbf{z}_i\}_{i=1,\dots,N_s}$

Calcul de la fonction de vraisemblance :  $p(\mathbf{Z}|w_k)$

initialisation :  $m = 0$  et  $\mathbf{Z}_0 = \mathbf{Z}$

**répété**

$m = m + 1$

Extraction du point de vraisemblance maximale :

$$\hat{\mathbf{z}}_m = \mathbf{Z}_k | k = \arg \max_k (p(\hat{\mathbf{Z}}_m | w_k = w_1))$$

Calcul de la liste de points associée :

$$\mathbf{L}^m = \{\mathbf{Z}_i \in \mathbf{Z} | d_c(\hat{\mathbf{z}}_m, \mathbf{z}_i) < \kappa\}$$

Mise à jour du vecteur de vraisemblance

$$\hat{\mathbf{Z}}_{m+1} = \overline{\hat{\mathbf{Z}}_m \cap \mathbf{L}^m}$$

**jusqu'à** critère d'arrêt atteint :  $p(\hat{\mathbf{Z}}_m | w_k = w_1) < \delta$

$M = m$

**retourne** le vecteur des points sélectionnés :

$$\hat{\mathbf{Z}} = \{\hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2, \dots, \hat{\mathbf{z}}_M\}$$


---

### 4.1.3 Fusion des détections issues des quatre nappes laser

Comme argumenté dans l'introduction de ce chapitre, nous proposons ici une méthode de détection issue de la combinaison des quatre nappes du capteur laser. L'idée principale est de combattre les fausses détections qui peuvent apparaître dans une nappe isolée en cherchant une confirmation de détection dans les trois autres. La figure 4.7 illustre ce problème en proposant la détection de trois piétons. Cet exemple montre que l'image de points dans chaque nappe est différente. Cette remarque s'explique par l'écart angulaire ( $1.07^\circ$ ) qui sépare chaque plan laser. L'information acquise de l'environnement est donc plus riche car elle permet d'obtenir un relevé télémétrique à différentes hauteurs. Une partie du problème de l'occultation est alors partiellement résolue comme, par exemple, celle d'un objet caché par un autre de plus petite hauteur.

#### Combinaison des quatre plans laser

Un piéton situé à une certaine distance du capteur doit théoriquement être coupé par un nombre connu  $N_l$  de nappes (cf. la figure 4.8). Pour une raison quelconque (couleur absorbante, transparence, occultation...), un piéton situé dans le champ de vue du capteur peut ne pas renvoyer d'« objets-piétons » dans un plan. On doit donc utiliser une méthode qui recherche un amas d'« objets-piétons » dans une zone de taille définie par l'imprécision de mesure du capteur.

Nous proposons alors d'utiliser une méthode non paramétrique, fondée sur une modélisation discrète de la densité de probabilité de chaque « objet-piéton » par la méthode des « fenêtres de Parzen » [Egammal *et al.*, 2000]. Dans un premier temps, toutes les positions  $N_d$  situées dans les quatre nappes laser  $\hat{Z} = \{\hat{z}_k\}_{k=1,\dots,N_d}$  sont projetées sur un même plan parallèle au sol noté  $I_d$ . La fusion des informations situées dans les quatre plans est alors effectuée par une fonction similaire au module d'extraction définie par :

$$p(\hat{Z}|w_k) = \frac{1}{N_l} \sum_{i=1}^{N_d} \psi(\hat{z}_k, \hat{z}_i) \quad (4.24)$$

Avec  $\psi(\hat{z}_k, \hat{z}_i)$  le noyau gaussien caractérisé par  $\Sigma_\psi$  la matrice de covariance associée à l'imprécision du capteur laser dans les deux dimensions ( $x$  et  $y$ ).

$$\Sigma_\psi = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \quad (4.25)$$

Comme le montre la figure 4.8, un piéton est coupé par un certain nombre de nappes  $N_l$  en fonction de sa distance avec le capteur. Les équations qui permettent de prendre en compte cette dernière remarque sont

$$\text{avec } N_l(i) = 1 \text{ si } (0 < (hc + D \cdot \tan(\phi_i)) < H) \quad (4.26)$$

$$N_l = \sum_{i=1}^4 N_l(i) \quad (4.27)$$

Ces expressions prennent en compte la hauteur  $H$  d'un piéton, la résolution angulaire entre les plans  $\phi_i$  en fonction de la distance  $D$  séparant l'obstacle du véhicule et de la hauteur du capteur par rapport au sol  $h_c$ .



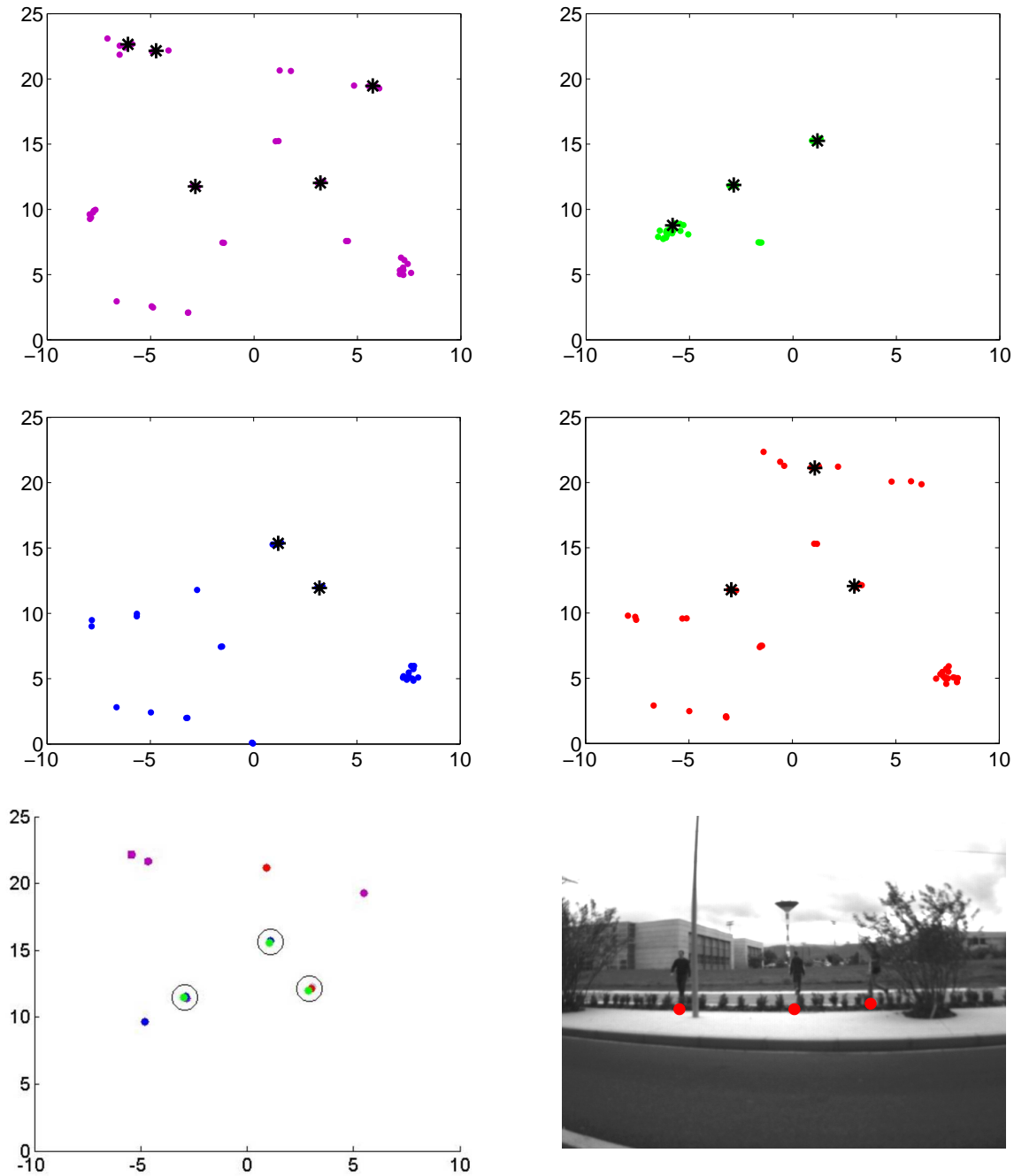


FIG. 4.7 – Les quatre images laser représentent la scène observée suivant les quatre plans laser du capteur. Les piétons détectés par le module d'extraction de piéton sont représentés par une étoile noire. Le résultat final de la fusion des quatre nappes est montré en bas à gauche par un cercle noir. Finalement, le résultat final de détection est projeté dans l'image vidéo qui permet de vérifier l'exactitude du résultat. Il est intéressant de remarquer la différence des informations contenues dans les quatre plans laser de la même scène observée. Le nombre de piétons détectés et leurs positions sont donc différents d'un plan à l'autre. La fusion des quatre plans permet (dans cet exemple), d'éliminer les fausses détections et de ne conserver que les trois piétons.

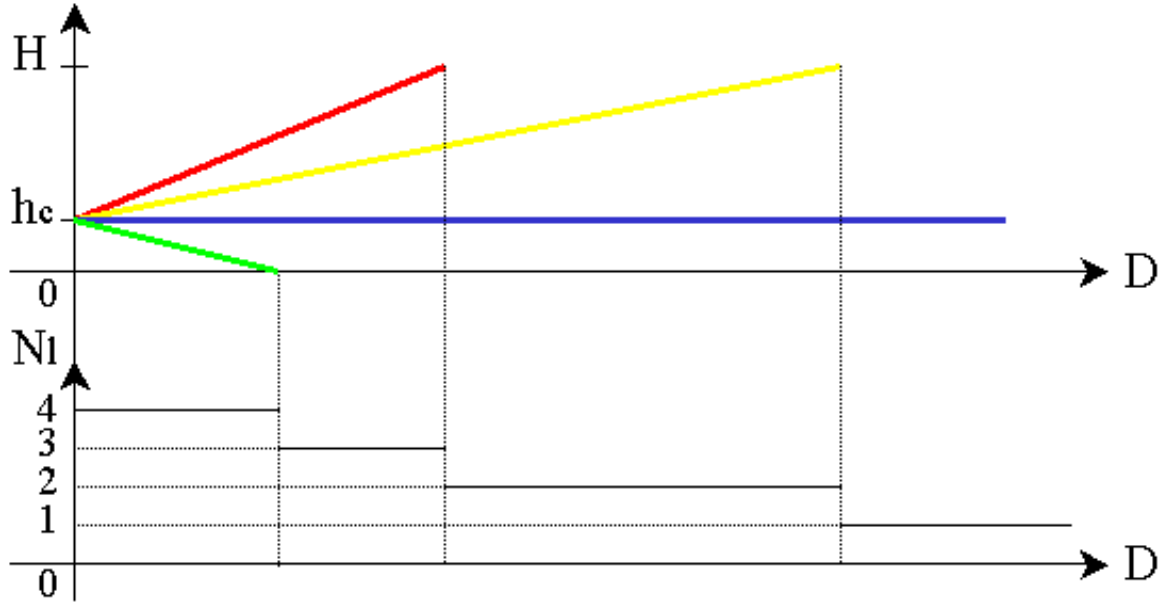


FIG. 4.8 – Nombre de nappes  $N_l$  en fonction de la distance objet/capteur.

### Localisation d'un piéton

Une fois la vraisemblance de la combinaison des quatre plans laser définie, il convient de rechercher la position de l'« objet-piéton » possédant la plus forte probabilité. La position dominante  $\tilde{\mathbf{z}}_m$  possédant le poids maximum associé à l'amas, notée  $\hat{\mathbf{Z}}_k|k$  est donnée par l'estimateur du maximum de vraisemblance comme position d'un piéton dans l'image :

$$\tilde{\mathbf{z}}_m = \tilde{\mathbf{Z}}_k|k = \arg \max_k (p(\hat{\mathbf{Z}}|w_k = w_1)) \quad (4.28)$$

### Localisation de plusieurs piétons

Dans le cas d'une image dans laquelle plusieurs zones sont présentes, la relation (4.21) fournit la position du centre de gravité de l'amas le plus représentatif. En considérant que la taille  $\beta$  de cet amas est grossièrement connue, une liste de points  $\hat{\mathbf{L}}^m$  (en considérant l'indice  $m$  associé au numéro de l'amas courant détecté) est éliminée :

$$\hat{\mathbf{L}}^m = \{N_d \in I_d | d_c(\tilde{\mathbf{z}}_m, \hat{\mathbf{z}}_i) < \beta\} \quad (4.29)$$

Le seuil  $\beta$  est fixé en choisissant la taille de la fenêtre de validation de telle sorte que la probabilité de trouver les mesures laser correctes soit égale à  $P_\beta$ .  $\beta$  est lié à  $P_\beta$  par la relation :

$$P_\beta = p(\chi_{nz}^2 \leq \beta) \quad (4.30)$$

où  $\chi_{nz}^2$  désigne une variable qui suit une loi du  $\chi^2$  à  $nz$  degré de liberté.

Une nouvelle liste de points candidats  $\hat{\mathbf{L}}^{m+1}$  est alors constituée. Les relations (4.24) et (4.28) sont réitérées selon l'algorithme 8 tant que l'estimateur du maximum de vraisemblance contient une valeur supérieure à un seuil de confiance  $v \in [0, 1]$ .

**Algorithme 8** Algorithme de détection d'« objets-piétons » issues des quatre nappes laser avec un estimateur non paramétrique à noyau

**Entrées :** Liste des centres de gravité situés dans les quatre nappes laser qui n'ont pas été filtrés par le module de filtrage des fausses détections :  $\hat{\mathbf{Z}} = \{\hat{\mathbf{z}}_i\}_{i=1,\dots,Nd}$

Calcul de la fonction de vraisemblance :  $p(\hat{\mathbf{Z}}|w_k)$

Initialisation :  $m = 0$  et  $\hat{\mathbf{Z}}_0 = \hat{\mathbf{Z}}$

**répété**

$m = m + 1$

Extraction du point de vraisemblance maximale :

$$\tilde{\mathbf{z}}_m = \hat{\mathbf{Z}}_k | k = \arg \max_k (p(\tilde{\mathbf{Z}}_m | w_k = w_1))$$

Calcul de la liste de points associée :

$$\hat{\mathbf{L}}^m = \{\hat{\mathbf{z}}_i \in \hat{\mathbf{Z}} | d_c(\tilde{\mathbf{z}}_m, \hat{\mathbf{z}}_i) < \beta\}$$

Mise à jour du vecteur de vraisemblance

$$\tilde{\mathbf{Z}}_{m+1} = \overline{\tilde{\mathbf{Z}}_m \cap \hat{\mathbf{L}}^m}$$

**jusqu'à** critère d'arrêt atteint :  $p(\tilde{\mathbf{Z}}_m | w_k = w_1) < v$

$M = m$

**retourne** le vecteur des points sélectionnés :

$$\tilde{\mathbf{Z}} = \{\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2, \dots, \tilde{\mathbf{z}}_M\}$$


---

#### 4.1.4 Association et suivi multi-objets

Dans cette section, nous décrivons la modélisation de notre système de suivi de piétons à partir de détections laser. Le rôle de cette étape est de produire une information temporellement plus cohérente que les mesures qui l'alimentent, mais aussi du nombre d'objets présents et de leurs identités respectives. Cette fonctionnalité peut permettre également de filtrer des observations peu probables car peu mesurées durant une période ou de maintenir une piste lorsque l'observation défaille pendant quelques périodes.

Le choix de l'algorithme de filtrage dépend donc directement de l'application visée. Dans le cas du suivi de piéton, nous supposons ne pas avoir de connaissance ou d'hypothèse *a priori* sur son mouvement difficilement prévisible (changement de rythme, inversion de direction, arrêt brutal, etc.). Les filtres particuliers permettent de surmonter cette difficulté parce qu'ils gèrent naturellement les non-linéarités de trajectoire. De nature probabiliste, cette modélisation facilite la prise en compte des différentes incertitudes apportées par l'inexactitude des modèles d'évolutions utilisés, de l'imprécision des capteurs, des différents bruits d'environnement, etc.

Un tel filtre permet donc de suivre une trajectoire à partir d'observations reçues séquentiellement. Mais, dans un cas réel, il est fréquent que plusieurs pistes existent en même temps, donnant naissance à des données leur étant propres. Dans ce cas, il faut choisir les données issues du capteur qui devront être utilisées par le filtre. C'est l'étape

d'association des données. De plus, un filtre ne gère pas naturellement l'initialisation et la destruction des pistes. Un module de gestion des pistes est donc utilisé (cf. la figure 4.9).

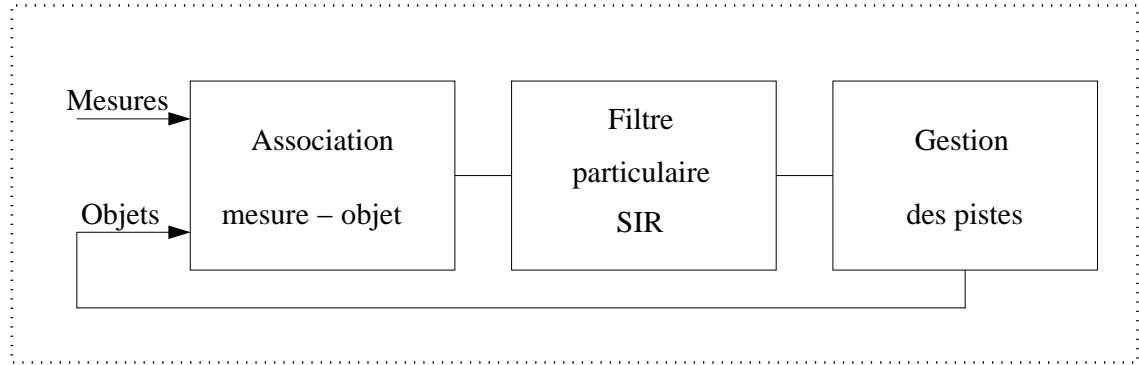


FIG. 4.9 – Algorithme de suivi de piéton.

### Filtre particulaire SIR

Classiquement, le filtre SIR (Sequential Importance Resampling) est formalisé par deux processus aléatoires à temps discret : le processus d'état  $\mathbf{X}_k$  (4.31) et le processus de mesure  $\mathbf{Z}_k$  (4.32).

$$\mathbf{X}_k = f(\mathbf{X}_{k-1}, \mathbf{w}_k) \quad (4.31)$$

$$\mathbf{Z}_k = h(\mathbf{X}_k, \mathbf{v}_k) \quad (4.32)$$

Aucune hypothèse de linéarité n'est faite sur les fonctions  $f$ ,  $h$ , et où  $\mathbf{w}_k$ ,  $\mathbf{v}_k$  sont des bruits blancs mutuellement indépendants pouvant être éventuellement non gaussiens.

### Modèle d'état de la configuration multi-piétons

Afin de suivre simultanément un nombre variable d'objets, l'état encode le nombre et la configuration des objets perceptibles dans le champ de mesure :

$$\mathbf{X}_k = (O_k, \mathbf{x}_k^i), \quad \forall i \in \{1, \dots, O_k\} \quad (4.33)$$

où  $O_k$  est le nombre d'objets suivis à l'instant  $k$  et chaque objet  $\mathbf{x}_k^i$  est défini par :

$$\mathbf{x}_k^i = (\mathbf{p}_k^i, \mathbf{v}_k^i, \mathbf{I}_k^i), \quad \forall i \in \{1, \dots, O_k\} \quad (4.34)$$

La position et l'incertitude relatives de l'objet  $i$  par rapport au véhicule à l'instant  $k$  sont définies par  $\mathbf{p}_k^i = (\mathbf{x}_k^i, \mathbf{y}_k^i, \sigma_{\mathbf{x}}^{k,i}, \sigma_{\mathbf{y}}^{k,i})$  où  $\mathbf{x}_k^i, \mathbf{y}_k^i$  et  $\sigma_{\mathbf{x}}^{k,i}, \sigma_{\mathbf{y}}^{k,i}$  sont respectivement la position du centre de gravité de l'objet et son incertitude associée. La vitesse relative de l'objet  $\mathbf{v}_k^i$  est définie par  $\mathbf{v}_k^i = (\dot{\mathbf{x}}_k^i, \dot{\mathbf{y}}_k^i, \dot{\sigma}_{\mathbf{x}}^{k,i}, \dot{\sigma}_{\mathbf{y}}^{k,i})$  où  $\dot{\mathbf{x}}_k^i, \dot{\mathbf{y}}_k^i$  et  $\dot{\sigma}_{\mathbf{x}}^{k,i}, \dot{\sigma}_{\mathbf{y}}^{k,i}$  sont respectivement la vitesse et son incertitude associée. Enfin,  $\mathbf{I}_k^i = (id_k^i, age_k^i)$  regroupe l'identification ( $id_k^i$ ) et l'âge ( $age_k^i$ ) de la piste  $i$  dans l'image de profondeur.

### Modèle d'évolution dynamique des piétons

Afin d'améliorer l'efficacité du suivi, nous limitons l'espace d'état à l'étendue que peuvent

physiquement atteindre les objets suivis. En effet, la performance d'un filtre dépend directement du modèle dynamique utilisé et de la capacité de ce dernier à décrire l'évolution de l'entité au cours du temps. Les résultats expérimentaux présentés en Annexe B illustrent et quantifient le gain du suivi en position apporté par un modèle d'évolution à mouvement circulaire comparé à un modèle d'évolution à vitesse constante. Nous avons donc choisi de décrire la dynamique d'un piéton par le modèle suivant :

$$F_k = \begin{bmatrix} 1 & \frac{\Delta T \sin(\theta_k)}{\theta_k} & 0 & -\frac{\Delta T}{\theta_k} \cdot (1 - \cos(\theta_k)) \\ 0 & \cos(\theta_k) & 0 & -\sin(\theta_k) \\ 0 & \frac{\Delta T}{\theta_k} \cdot (1 - \cos(\theta_k)) & 1 & -\frac{\Delta T \sin(\theta_k)}{\theta_k} \\ 0 & \sin(\theta_k) & 0 & \cos(\theta_k) \end{bmatrix} \quad (4.35)$$

où  $\theta_{k+1} = \theta_k + b_g$  avec  $b_g \sim \mathcal{N}(0, \sigma_g)$ .  $\sigma_g$  est l'écart type du bruit sur le cap de la trajectoire du piéton.  $\Delta T$  est la différence de temps entre deux instants successifs  $k$  et  $k - 1$ .

### Association mesure-objet

Un algorithme de suivi multi-objets fondé sur un filtre à particules comporte généralement une étape d'association de données. L'efficacité d'une telle méthode dépend directement des observations disponibles et des informations connues *a priori*. Dans le cas où une seule mesure est sélectionnée, les algorithmes de filtrage présentés dans le chapitre 3 peuvent être appliqués directement. Dans le cas contraire, une modélisation probabiliste des différentes associations possibles peut être construite.

Dans le cas du suivi multi-piétons à partir d'un unique capteur laser, deux hypothèses d'association sont considérées :

- ▷ **H1** : une mesure peut provenir d'un objet ou être une fausse alarme. Cela induit qu'une mesure peut être délivrée par un objet effectivement suivi ou par une perturbation de l'observation, notamment lors d'une mauvaise segmentation de l'image de profondeur,
- ▷ **H2** : à un instant donné, un objet peut être à l'origine d'aucune ou d'au plus une mesure. Cette hypothèse a beaucoup de sens dans le cadre applicatif de la perception par LiDAR. En effet, un piéton qui n'est à l'origine d'aucune mesure peut correspondre à une situation d'occultation. D'autre part, un piéton ne peut pas correspondre à deux parties différentes de l'image laser au même moment.

Nous proposons de construire une matrice  $\mathbf{A}$ , appelée matrice d'associations, qui représente les associations possibles entre les objets prédits et les mesures provenant du module de détection (cf. l'algorithme 7). Les notations suivantes sont alors utilisées :

- ▷ on désigne par  $\tilde{\mathbf{X}}_k = \{\tilde{\mathbf{x}}_k^i\}_{i=1,\dots,O}$  l'ensemble des positions  $O$  prédites des objets suivis à l'instant  $k$ ,
- ▷ on désigne par  $\mathbf{Z}_k = \{\mathbf{z}_k^j\}_{j=1,\dots,m_k}$  l'ensemble des mesures  $m_k$  reçues à l'instant  $k$ .

Au temps  $k$ , la matrice d'associations  $\mathbf{A}$  représente graphiquement toutes les possibilités d'associations mesure-objet possibles. Dans cette représentation, les  $O$  objets représentent le nombre de lignes et les  $m_k$  mesures représentent le nombre de colonnes de la matrice  $\mathbf{A}$ . La figure 4.10 illustre un exemple de matrice  $\mathbf{A}$ .

Dans le cadre de l'utilisation d'un filtre à particules, nous proposons de calculer la probabilité de l'association  $p(\mathbf{z}_k^j | \tilde{\mathbf{x}}_k^i)$  entre l'objet  $i$  et la mesure  $j$  par une approche d'association non paramétrique. L'idée est d'utiliser lors de l'association mesure-objet, le jeu

$$\begin{array}{c}
 \mathbf{A} = \begin{array}{c} \text{O lignes} \\ \mathbf{d}_{ij} \end{array} \begin{array}{c} \mathbf{m}_k \text{ colonnes} \\ \left[ \begin{array}{cccc} \mathbf{d}_{11} & \mathbf{d}_{12} & \mathbf{d}_{13} & \cdots \\ \mathbf{d}_{21} & \mathbf{d}_{22} & \mathbf{d}_{23} & \cdots \\ \mathbf{d}_{31} & \mathbf{d}_{32} & \mathbf{d}_{33} & \cdots \\ \vdots & \vdots & \vdots & \mathbf{d}_{0m_k} \end{array} \right] \end{array}
 \end{array}$$

FIG. 4.10 – Exemple d’une matrice d’associations.

de particules dans sa globalité sans perdre le caractère quelconque de la dispersion spatiale du nuage par une approximation gaussienne par exemple [Gidel *et al.*, 2008b].

En effet, les méthodes non paramétriques permettent de prendre en compte les échantillons et leur répartition spatiale dans l’espace des paramètres. La conséquence est une estimation de la distribution  $p(\mathbf{Z}_{1:k}|\tilde{\mathbf{X}}_k)$  plus proche de la réalité. En effet, la probabilité  $P$  que le vecteur  $\mathbf{X}$  soit dans la région  $R$  est donnée par  $P = \int_R p(\mathbf{X}')d\mathbf{X}'$ . Si  $R$  est petit et que  $P(\mathbf{X})$  est continu, alors  $\int_R p(\mathbf{X}')d\mathbf{X}' = P(\mathbf{X})$ . Finalement, un estimé de  $P(\mathbf{X})$  peut alors être donné par  $P(\mathbf{X}) = \frac{k/n}{V}$  avec  $k$  le nombre d’échantillons qui sont dans  $R$ ,  $V$  le volume délimité par  $R$  et  $n$  le nombre total d’échantillons.

Il existe deux techniques qui permettent d’engendrer une suite de régions qui satisfont de bonnes conditions d’estimation :

- ▷ en fixant le volume de la région comme une fonction de  $n$ , par exemple :  $V_n = \frac{1}{\sqrt{n}}$ .

**C’est la méthode des « fenêtres de Parzen » (KDE).**

- ▷ en adaptant la taille des régions au nombre d’échantillons  $k_n$  fixé en fonction de  $n$ , par exemple :  $k_n = \sqrt{n}$

**C’est la méthode des K plus proches voisins (KNN).**

Dans le cadre de nos travaux, la méthode des « fenêtres de Parzen » a été choisie car elle utilise l’estimateur par noyau (Kernel Density Estimation) qui permet de limiter l’influence des données sur une zone définie par le noyau  $\Sigma_A$ . Le calcul de la probabilité d’association entre une observation  $\mathbf{z}_k^j$  et l’ensemble de ces voisins (particules)  $\tilde{\mathbf{x}}_k^i$  est alors définie par :

$$d_{ij} = p(\mathbf{z}_k^j|\tilde{\mathbf{x}}_k^i) = \frac{1}{N} \sum_{n=1}^N \phi(\mathbf{z}_k^j, \mathbf{x}_k^{i,n}) \quad (4.36)$$

avec  $\phi(\mathbf{z}_k^j, \mathbf{x}_k^{i,n})$  le noyau qui permet de modifier la zone d’influence d’un point par rapport aux autres. Nous avons fait le choix classique d’utiliser un noyau de type gaussien défini par :

$$\phi(\mathbf{z}_k^j, \mathbf{x}_k^{i,n}) = \frac{1}{2\pi|\Sigma_A|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{z}_k^j - \mathbf{x}_k^{i,n})\Sigma_A^{i(-1)}(\mathbf{z}_k^j - \mathbf{x}_k^{i,n})\right] \quad (4.37)$$

avec  $\Sigma_A^i = \Sigma_S^i + \Sigma_\psi$  où  $\Sigma_S^i$  est la matrice de covariance de l'objet  $i$  suivi et  $\Sigma_\psi$  la matrice de covariance associée à l'imprécision de mesure du capteur laser.

### Fonction discriminante

Une fois la matrice  $\mathbf{A}$  complètement calculée, les associations mesure-objet pour lesquelles les probabilités sont en-dessous d'un seuil empirique  $C$  sont considérées comme inattendues. Dans ce cas  $d_{ij}$  est mis à 0.

$$d_{ij} = \begin{cases} p(\mathbf{z}_k^j | \tilde{\mathbf{x}}_k^i) > C \\ 0 & \text{sinon} \end{cases} \quad (4.38)$$

### Localisation d'une association mesure-objet

Une fois la matrice d'associations  $\mathbf{A}$  définie et filtrée par la fonction discriminante, il convient de sélectionner la position de l'association mesure-objet possédant la plus forte probabilité. L'association dominante notée  $\mathbf{Z}_k^{j*}$  est donnée par la mesure qui possède le poids maximum à l'itération  $ii$ .

$$\mathbf{Z}_k^{j*} | ii = \arg \max_{ii} (p(\mathbf{z}_k^j | \tilde{\mathbf{x}}_k^i)) \quad (4.39)$$

### Localisation de plusieurs associations mesure-objet

Dans le cas d'une matrice  $\mathbf{A}$  dans laquelle plusieurs coefficients  $d_{ij}$  sont supérieurs à zéro, la relation (4.39) fournit la position de l'association mesure-objet la plus représentative. En considérant l'hypothèse **H2**, la probabilité  $\{d_{ij}\}$  associant l'objet  $i = q$  avec l'observation  $j = t$  est éliminée, ainsi que toute la ligne  $i = q$  et toute la colonne  $j = q$ .

$$\mathbf{A}_k^{ii+1} = \mathbf{A}_k^{ii} - \mathbf{I} \cdot \{d_{ij}\}_{i=q \cap 1 \leq i \leq O}^{j=t \cap 1 \leq j \leq m_k} \quad (4.40)$$

Une nouvelle matrice d'associations  $\mathbf{A}_k^{ii+1}$  est alors constituée. La relation (4.39) est réitérée tant que l'estimateur du maximum de vraisemblance contient une valeur supérieure à zéro.

### Exemple d'évolution d'une matrice d'associations $\mathbf{A}$ à l'instant $k$

Au temps  $k$ , soit  $\tilde{\mathbf{X}}_k = \{\tilde{\mathbf{x}}_k^i\}_{i=1,\dots,4}$ , les quatre positions d'objets prédits et  $\mathbf{Z}_k = \{\mathbf{z}_k^j\}_{j=1,\dots,4}$ , les quatre observations validées. En se limitant aux hypothèses **H1** et **H2**, on calcule les coefficients  $\{d_{ij}\}$  de la matrice d'associations  $\mathbf{A}$  de taille  $(i, j)$ . Après avoir tenu compte de la fonction discriminante, il s'est avéré que (cf. la figure 4.11) :

- ▷ l'observation numéro 1 peut être mise en relation avec trois objets  $d_{11}$ ,  $d_{31}$  et  $d_{41}$ .
- ▷ l'observation numéro 2 peut être mise en relation avec trois objets  $d_{12}$ ,  $d_{22}$  et  $d_{32}$ .
- ▷ l'observation numéro 3 peut être mise en relation avec trois objets  $d_{23}$ ,  $d_{33}$  et  $d_{43}$ .
- ▷ l'observation numéro 4 peut être mise en relation avec deux objets  $d_{14}$  et  $d_{44}$ .

### Gestion des pistes

Les pistes qui sont confirmées aux itérations précédentes sont suivies par notre module de type SIR/KDE qui a l'avantage de conserver le caractère non-gaussien de la distribution.

**H1** et **H2** permettent de traiter la matrice d'associations  $\mathbf{A}_k$  à l'instant  $k$  mais ce système ne permet pas de gérer les cas d'occultation ou autres des pistes.



$$\begin{aligned}
 \mathbf{A}_{k, ii} &= \begin{bmatrix} d_{11} & d_{12} & 0 & d_{14} \\ 0 & d_{22} & d_{23} & 0 \\ d_{31} & d_{32} & d_{33} & 0 \\ d_{41} & 0 & d_{43} & d_{44} \end{bmatrix} \\
 \mathbf{A}_{k, ii+1} &= \begin{bmatrix} 0 & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & 0 & d_{43} \end{bmatrix} \\
 \mathbf{A}_{k, ii+2} &= \begin{bmatrix} 0 & d_{23} \\ d_{41} & d_{43} \end{bmatrix} \quad \mathbf{A}_{k, ii+3} = \begin{bmatrix} 0 \end{bmatrix}
 \end{aligned}$$

FIG. 4.11 – Exemple d'évolution d'une matrice d'associations  $\mathbf{A}_k$  au temps  $k$ . Dans le cas où plusieurs associations sont possible (une colonne de  $\mathbf{A}_k$  contient plusieurs  $d_{ij}$ ), on commence par associer le couple mesure-objet qui possède la plus forte probabilité d'association  $d_{ij}$ . Dans le cas de notre espace joint (hypotheses **H1** et **H2**), la distance  $d_{ij}$  possédant la plus forte probabilité est  $d_{14}$ . Le couple associant l'objet numéro 1 avec l'observation numéro 4 est alors éliminé avec sa ligne et sa colonne complète afin de supprimer toute correspondance future d'une mesure avec d'autres objets. L'algorithme est réitéré tant qu'une observation reste à associer. On remarque que dans notre exemple l'observation numéro 1 ne trouve finalement pas d'objet avec qui s'associer. Dans ce cas, le module de gestion des pistes décide d'en créer une nouvelle.

Afin de résoudre ce problème, un système de calcul de probabilité est mis en place. Si lors d'une itération, une piste ne trouve aucune mesure avec qui s'associer, alors la probabilité d'existence de la piste est diminuée d'une certaine quantité fixe  $\delta_p$ . Bien évidemment, les pistes  $\mathbf{x}_k^i$  qui ont une probabilité inférieure à « 1 » voient celle-ci augmenter de la même quantité  $\delta_p$  si elles sont associées à une mesure.

Afin de modifier le nombre d'objets suivis, le module de gestion des pistes est défini de la manière suivante :

- ▷ si une mesure ne peut être associée à aucune piste (une colonne de  $\mathbf{A}$  contient que des zéros), alors le module de gestion va proposer un nouvel objet d'une probabilité d'existence  $\delta_p$ ,
- ▷ si une piste a une probabilité d'existence inférieure à un seuil  $P_{min}$ , alors le module de gestion va proposer sa destruction,
- ▷ si une piste a une probabilité d'existence supérieure à un seuil  $P_{max}$ , alors le module de gestion va la confirmer. Sa probabilité d'existence est alors fixée à 1.

Les différentes étapes de ce filtre sont données dans l'algorithme 9.

---

**Algorithme 9** Association et suivi temporel : SIR / KDE.
 

---

1. Pour  $k = 0$ , générer  $N$  échantillons à partir de chaque mesure  $j = 1, \dots, m_k$ , i.e  $\{\mathbf{x}_0^{i,n}\}_{n=1}^N = \{\mathbf{x}_0^{1,n}, \dots, \mathbf{x}_0^{m_k,n}\}_{n=1}^N$  où  $\mathbf{x}_0^{i,n} = p(\mathbf{X}_0^j)$ .
  2. Calculer la matrice  $\mathbf{A}_k$  pour toutes les observations  $m_k$  et tous les objets  $O$  si ( $d_{ij} \geq C$ )  
 $\mathbf{A}_k = d_{ij}$  avec  $1 \leq i \leq 0$  et  $1 \leq j \leq m_k$  où  $d_{ij}$  est la probabilité d'association entre l'objet  $i$  et l'observation  $j$  en utilisant les  $N$  particules suivant l'équation (4.36) alors  $\{\mathbf{x}_k^{i,n}\} = \{\tilde{\mathbf{x}}_k^{i,n}\}$ .  
 sinon  
 $\mathbf{A}_k^{i,j} = 0$  alors  $\{\mathbf{x}_k^{i,n}\} = \{\mathbf{x}_{k-1}^{i,n}\}$  aller à l'item 5.
  3. Calcul des poids  $\pi_k^i = p(\mathbf{z}_k | \mathbf{x}_k)$  puis normalisation, i.e,  $\pi_k = \frac{\pi_k}{\sum_{l=1}^N \pi_k}$ .
  4. Générer un nouveau nuage de particules  $\{\mathbf{x}_k^{i,n*}\}_{n=1}^N$  après rééchantillonnage de  $\{\mathbf{x}_k^{i,n}\}_{n=1}^N$ , selon la probabilité  $Pr(\mathbf{x}_k^{i,n*} = \mathbf{x}_k^{i,n}) = \pi_k^n$ .
  5. Prédire les nouvelles particules,  $\mathbf{x}_{k+1}^{i,n} = f(\mathbf{x}_k^{i,n*}, \mathbf{v}_k)$ ,  $n = 1, \dots, N$  en bruitant chaque particule à partir du modèle d'évolution.
  6.  $k = k + 1$  aller à l'item 2.
- 

## 4.2 Fusion multisensorielle laser/vidéo

### 4.2.1 Introduction

La fusion de données est un domaine qui permet de mélanger des informations provenant de différentes sources afin d'extraire une estimation ou une décision plus précise et plus robuste.

La première partie de ce chapitre décrit une méthode de détection et de suivi de piéton à partir d'un unique capteur laser quatre plans. Dans un environnement urbain complexe tels que ceux étudiés dans le projet LOVE, les performances d'un tel système de détection

se trouvent ainsi limitées par la nature même de l'information délivrée. En effet, son incapacité à délivrer des informations sur la hauteur, le contour ou sur la couleur des entités détectées, complique notamment la tâche de classification de ces objets en tant que piéton ou non. Par exemple, de nombreuses fausses détections sont dues à une mauvaise classification entre un piéton et un arbre ou entre un piéton et un panneau de la circulation qui sont des objets ayant une forme géométrique similaire dans une image de profondeur.

La fusion multisensorielle laser/vidéo vise donc à pallier aux limitations d'un capteur LiDAR comme celui utilisé pour le projet LOVE. Une étude sur la détection de piéton multisensorielle présentée par [Gavrila *et al.*, 2001], indique que la combinaison d'informations provenant d'un capteur laser avec une caméra vidéo semble être une bonne solution à développer de par la nature même de ces deux capteurs.

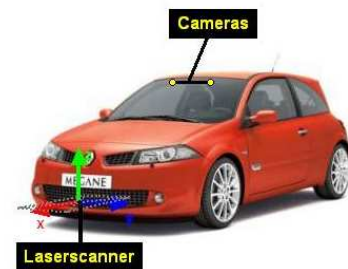
En effet, par sa capacité naturelle à saisir des informations de forme, de texture et de couleur des objets, de nombreuses méthodes de détection des piétons ont été développées à partir de caméras vidéo dans l'objectif de se rapprocher le plus possible des performances du système de perception humain qui est certainement le meilleur système actuel de détection. Pour une large revue de ces différentes méthodes existantes, le lecteur pourra consulter [Gavrila, 1999] où sont énumérées des approches fondées 2D et 3D.

Cependant, la forte sensibilité aux conditions lumineuses et l'impossibilité d'obtenir directement l'information de profondeur avec précision ont entre autres raisons, fait naître un intérêt pour la mise au point d'une méthode de détection à partir de capteurs actifs comme le radar ou le laser.

Cette section présente une architecture de détection et de suivi de piéton à partir d'une combinaison d'informations fournie par un capteur laser et une caméra vidéo. Après une rapide description de la plateforme utilisée, nous présentons la stratégie de fusion développée.

### 4.2.2 Description et positionnement des capteurs sur le véhicule

Le véhicule VELAC du LASMEA est équipé d'un système d'acquisition multisensorielle composé d'un télémètre quatre plans IBEO LDML positionné au centre du pare-chocs et d'une caméra en niveau de gris placée au niveau du rétroviseur intérieur. Le véhicule de Renault est équipé de manière similaire avec un télémètre quatre plans IBEO Alasca au centre du pare-chocs et de deux caméras SMAL en niveau de gris placées de part et d'autre du rétroviseur intérieur.



### 4.2.3 Contributions et algorithmes proposés

On a vu au chapitre 3 que la fusion multisensorielle revient à résoudre un problème de combinaison de plusieurs sources de données afin d'obtenir une nouvelle information de meilleure qualité. Une méthode a ainsi été développée dans ce but.

Cette méthode est fondée sur une architecture centralisée qui consiste à fusionner les mesures délivrées par les modules de détection basés laser et vidéo.

Cet algorithme (cf. la figure 4.12) s'appuie sur le même formalisme que l'algorithme 9 qui détaille notre méthode de suivi de piéton. En effet, le moteur de suivi repose sur un

filtre particulière SIR, le modèle d'évolution utilise le modèle à mouvement circulaire et enfin l'association des données est fondée sur les méthodes non paramétriques.

Les principales contributions de cet algorithme sont premièrement, la modélisation de la fonction de vraisemblance par le mélange d'une loi gaussienne avec une loi uniforme afin de mieux prendre en compte toute l'information disponible à chaque itération  $k$  et deuxièmement, le calcul d'une note de fusion qui prend en compte la qualité des détections (ou pistes) renseignées par les algorithmes situés en amont [Gidel *et al.*, 2009b].

### Formulation du système de fusion centralisée

#### Vecteur d'entrée et de sortie du module de fusion

Ce paragraphe décrit les entrées et sorties du module de fusion multi-capteurs permettant de détecter et de suivre les piétons à partir de mesures émanant d'un algorithme de détection vidéo et d'un algorithme de détection laser. Dans le cadre de nos travaux, une note de détection globale ( $NDG$ ) est associée aux sorties des modules mono-capteur afin de renseigner sur la confiance qu'un algorithme de détection accorde à la classification d'un objet comme piéton ou non. Le vecteur d'entrée du module de fusion pour chaque piéton  $j$  peut donc s'écrire de la manière suivante :

$$\mathbf{Z}_k = \{\mathbf{z}_k^j\}_{j=1,\dots,m_k} \quad (4.41)$$

avec chaque piéton détecté  $\mathbf{z}_k^j$  défini par :

$$\mathbf{z}_k^j = (\mathbf{p}_k^j, \gamma_k^j)_{j=1,\dots,m_k} \quad (4.42)$$

La position, les dimensions et les incertitudes associées à la détection  $j$  sont définies par  $\mathbf{p}_k^j$ . La note de détection globale  $\gamma_k^j$  évaluant la fiabilité de la détection est une probabilité telle que  $\gamma_k^j \in ]0, 1]$ .

De plus, le projet demande que les sorties du module de fusion fournissent pour chaque piéton  $i$  suivi un facteur de confiance sur la fusion des données ( $FCFD$ ), exprimant la confiance attribuée au piéton détecté par l'algorithme. Le vecteur de sortie du module de fusion pour chaque piéton  $j$  peut s'écrire :

$$\mathbf{X}_k = \{\mathbf{x}_k^i\}_{i=1,\dots,O} \quad (4.43)$$

avec chaque piéton suivi  $\mathbf{x}_k^i$  défini par :

$$\mathbf{x}_k^i = (\mathbf{p}_k^i, \mathbf{v}_k^i, \mathbf{I}_k^i, \alpha_k^i)_{i=1, \dots, O} \quad (4.44)$$

La position, les dimensions et les incertitudes associées à l'objet  $i$  à l'instant  $k$  sont définis par  $\mathbf{p}_k^i$ . La vitesse de l'objet et son identification sont données par respectivement  $\mathbf{v}_k^i$  et  $\mathbf{I}_k^i$ . Enfin le facteur de confiance sur la fusion des données  $\alpha_k^i$  est une probabilité telle que  $\alpha_k^i \in ]0, 1]$ .

#### Prise en compte de l'asynchronisme

En général, les observations délivrées par des capteurs de natures différentes sont asynchrones et délivrent donc rarement une même configuration de la scène observée. Pour effectuer une fusion correcte, il est nécessaire de déterminer à quel moment ont été effectuées les différentes mesures. L'information temporelle des différents capteurs doit donc être ramenée dans un même référentiel de temps. Dans le cas de mesures asynchrones, la

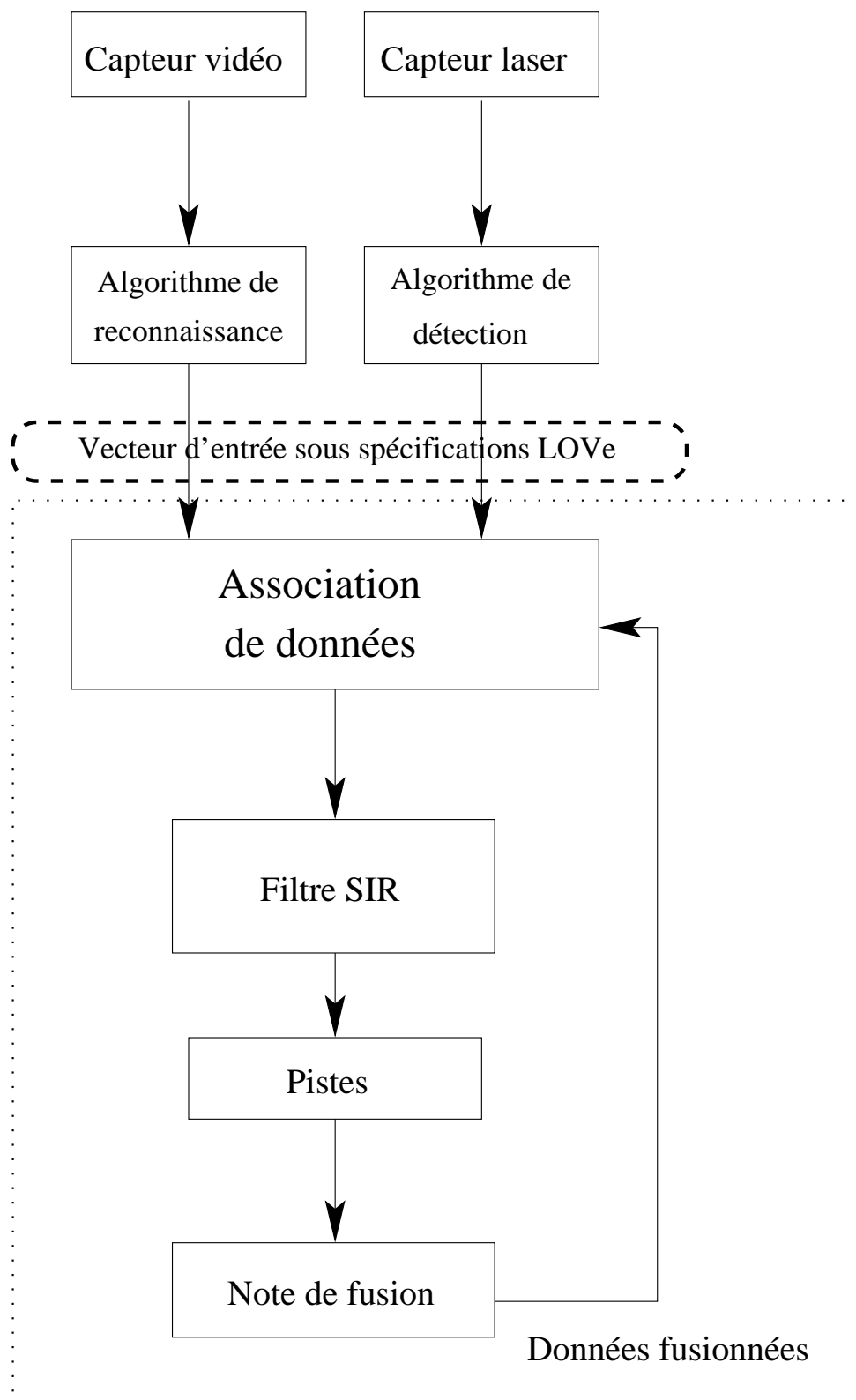


FIG. 4.12 – Architecture de fusion centralisée fondée sur un filtre particulaire SIR et une association non paramétrique des données. L'algorithme fusionne des mesures provenant d'algorithmes de détection et/ou de reconnaissance laser et vidéo.

correction de l'état peut être effectuée soit après chaque nouvelle observation, soit après avoir accumulé un nombre donné d'observations. Dans notre cas les données sont datées par les capteurs mais il faut encore tenir compte de la durée de traitement des algorithmes de détection qui retardent la livraison des mesures. Il faut donc commencer par enregistrer un certain nombre de données avant d'essayer de ramener dans un repère commun des observations mesurées à des instants différents. Deux possibilités s'offrent à nous pour recaler ces mesures dans un même repère.

La première solution consiste à synchroniser les mesures d'un capteur sur l'autre. L'échelle commune choisie est alors souvent celle du capteur le plus rapide. La méthode consiste à calculer une estimée de l'observation délivrée par l'algorithme de détection laser à l'instant immédiatement supérieur où l'algorithme vidéo retourne une observation de la scène (cf. la figure 4.13).

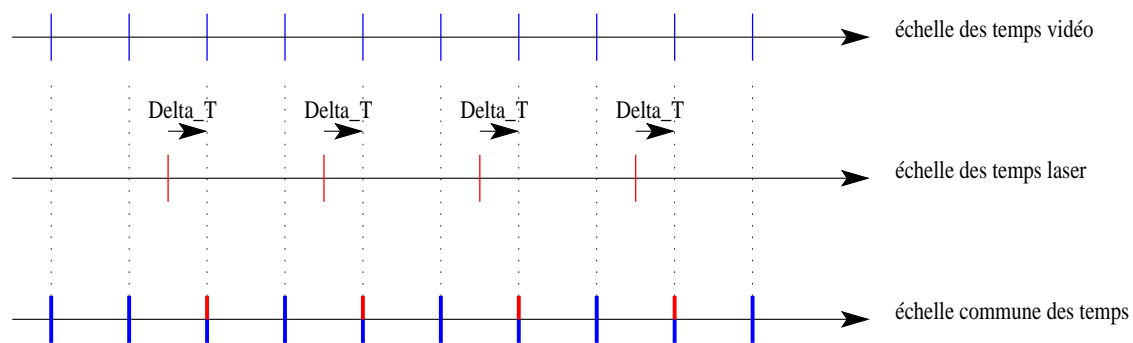


FIG. 4.13 – Description de la synchronisation des données d'un capteur sur l'autre dans une échelle commune.

La seconde solution consiste à ne pas synchroniser les mesures d'un capteur sur l'autre mais de garder toutes les dates disponibles. L'échelle commune comporte alors les dates des différents capteurs. La méthode consiste à trier les différentes données par ordre croissant de datation des différents capteurs (cf. la figure 4.14).

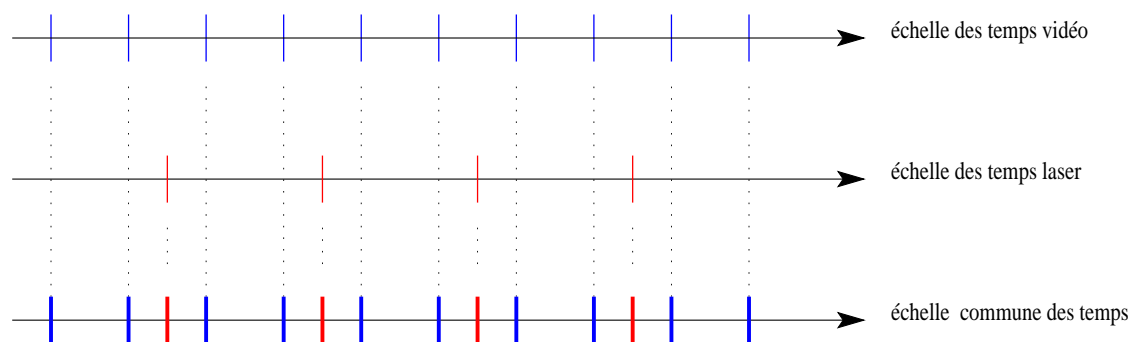


FIG. 4.14 – Description de la non synchronisation des données d'un capteur sur l'autre dans une échelle commune.

Nous avons choisi la deuxième approche car elle permet de conserver les hypothèses H1 et H2 (cf. le chapitre 3) sans ajouter l'hypothèse H3 qui propose qu'à un instant donné, un objet peut être à l'origine d'aucune ou de plusieurs mesures (délivrées par plusieurs capteurs).

### Association non paramétrique de données

Tout comme dans le cas de l'association multi-objets à partir d'un seul capteur (cf. § 4.1.4), il est important de construire une matrice d'associations  $\mathbf{A}$ , qui représente les associations possibles entre les objets et les mesures. Les notations suivantes sont alors utilisées :

- ▷ on désigne par  $\tilde{\mathbf{X}}_k = \{\tilde{\mathbf{x}}_k^i\}_{i=1,\dots,O}$  l'ensemble des positions prédites des objets suivis à l'instant  $k$ .
- ▷ on désigne par  $\mathbf{Z}_k = \{\mathbf{z}_k^j\}_{j=1,\dots,m_k}$  l'ensemble des mesures reçues à l'instant  $k$ .

Au temps  $k$ , la matrice d'associations  $\mathbf{A}$  représente graphiquement toutes les possibilités d'associations mesure-objet possibles. Dans cette représentation, les  $O$  hypothèses représentent le nombre de lignes et les  $m_k$  mesures représentent le nombre de colonnes de la matrice  $\mathbf{A}$  (cf. la figure 4.10).

Nous proposons de calculer la probabilité d'association  $p(\mathbf{z}_k^j | \tilde{\mathbf{x}}_k^i)$ , entre l'objet  $i$  et la mesure  $j$  par une approche d'association non paramétrique. L'idée principale de cette approche est d'utiliser lors de l'association mesure-objet, toute l'information disponible à l'instant  $k$ , c'est-à-dire, le jeu de particules dans sa globalité couplé à son  $NDG$  associé.

Nous introduisons une variable de Bernouilli  $w_h \in \{w_1, w_2\}$  pouvant prendre les deux états :  $w_h = w_1$  si l'évènement associé est classifié comme donnée fusionnée et  $w_h = w_2$  dans les autres cas. Cette méthode est non-supervisée car aucune connaissance *a priori* est nécessaire. La fonction de vraisemblance  $p(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^i | w_h)$  représente la probabilité qu'une particule 2D appartienne à une donnée fusionnée. Nous proposons de définir la fonction de vraisemblance  $p(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^i | w_h)$  avec un estimateur de densité fondé sur les « fenêtres de Parzen » qui permet de calculer la distance entre une observation  $\mathbf{z}_k^j$  et tous ses voisins  $\mathbf{x}_k^{i,n}$  telle que :

$$d_{ij} = p(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^i | w_h) = \frac{1}{N} \sum_{n=1}^N \psi(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^{i,n}) \quad (4.45)$$

où  $\psi(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^{i,n})$  est la fonction noyau qui permet de modifier la zone d'influence d'une observation avec ses voisins. Cette fonction noyau est composée d'un mélange de distributions gaussienne et uniforme (cf. la figure 4.15) afin de fusionner dans une même distribution toute l'information disponible fournie par les algorithmes mono-capteur en particulier l'information sur la fiabilité de la détection  $\gamma_k^j$ .  $\psi(\mathbf{z}_k^j, \mathbf{x}_k^{i,l})$  est décrite par :

$$\psi(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^{i,n}) = (1 - \gamma_k^j) \cdot \mathcal{U}\left(\frac{-1}{\gamma_k^j}, \frac{1}{\gamma_k^j}\right) + \gamma_k^j \cdot \exp[-\lambda_c \cdot d_c(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^{i,n})] \quad (4.46)$$

Le paramètre  $\lambda_c$  permet d'ajuster la valeur des poids. La distance  $d_c$  est une distance de Mahalanobis définie par :

$$d_c(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^{i,n}) = (\mathbf{z}_k^j - \tilde{\mathbf{x}}_k^{i,n}) \Sigma_\psi^{-1} (\mathbf{z}_k^j - \tilde{\mathbf{x}}_k^{i,n})^T \quad (4.47)$$

$$\Sigma_\psi = \Sigma_F + \Sigma_{SP2} \quad (4.48)$$

où  $\Sigma_F$  est la matrice de covariance donnée par l'algorithme de fusion centralisée et  $\Sigma_{SP2}$  est la matrice de covariance de l'objet donné par le module de détection mono-capteur.

La particule 2D qui obtient la plus forte probabilité  $\tilde{\mathbf{x}}_k^{i,l} \in w_1$  est choisie par l'estimateur du maximum de vraisemblance :

$$\mathbf{H}_k^{i,j} | (i, j) = \arg \max_{(i,j)} (p(\mathbf{z}_k^j | \tilde{\mathbf{x}}_k^i) | w_1) \quad (4.49)$$



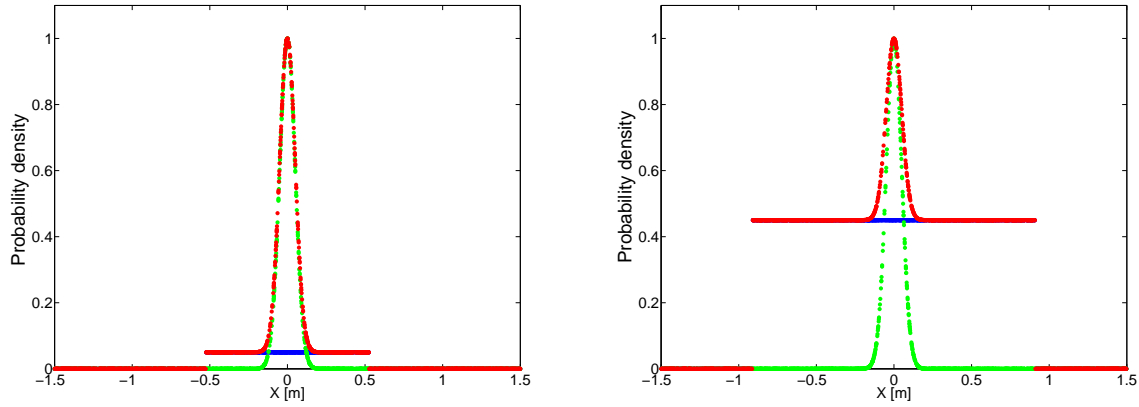


FIG. 4.15 – Un exemple d'un mélange de distributions gaussienne et uniforme, avec en bleu la distribution uniforme, en vert la distribution gaussienne et en rouge la distribution mélangée. A gauche,  $\sigma = 0.15$  m et  $NDG = 0.95$ . A droite,  $\sigma = 0.15$  m et  $NDG = 0.55$ .

### Fonction discriminante

Une fois la matrice  $\mathbf{A}$  calculée, les associations mesure-objet fusionnées pour lesquelles les probabilités sont en dessous d'un seuil empirique  $C$  sont considérées comme inattendues. Dans ce cas  $d_{ij}$  est mis à 0.

$$d_{ij} = \begin{cases} p(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^i | w_h) > C \\ 0 & \text{sinon} \end{cases} \quad (4.50)$$

### Localisation d'une association mesure-objet fusionnée

Une fois la matrice d'associations  $\mathbf{A}$  définie et filtrée par la fonction discriminante, il convient de sélectionner la position de l'association mesure-objet fusionnée possédant la plus forte probabilité. L'association dominante notée  $\mathbf{Z}_k^{j*}$  est donnée par la mesure qui possède le poids maximum à l'itération  $ii$  (cf. la figure 4.11).

$$\mathbf{Z}_k^{j*} | ii = \arg \max_{ii} (p(\mathbf{z}_k^j, \tilde{\mathbf{x}}_k^i | w_h)) \quad (4.51)$$

### Localisation de plusieurs associations mesure-objet fusionnées

Dans le cas d'une matrice  $\mathbf{A}$  dans laquelle plusieurs coefficients  $d_{ij}$  sont supérieurs à zéro, la relation (4.51) fournit la position de l'association mesure-objet la plus représentative. En considérant l'hypothèse **H2**, le couple  $\{d_{ij}\}$  associant l'objet  $i = q$  avec l'observation  $j = t$  sont éliminés, ainsi que toute la ligne  $i = q$  et toute la colonne  $j = q$ .

$$\mathbf{A}_k^{ii+1} = \mathbf{A}_k^{ii} - \mathbf{I} \cdot \{d_{ij}\}_{i=q \cap 1 \leq j \leq m_k}^{j=t \cap 1 \leq i \leq o} \quad (4.52)$$

Une nouvelle matrice d'associations  $\mathbf{A}_k^{ii+1}$  est alors constituée. La relation (4.51) est réitérée tant que l'estimateur du maximum de vraisemblance contient une valeur supérieure à zéro.

### Calcul de la vraisemblance

Un filtre à particules de type SIR permet d'approximer la distribution *a posteriori*  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$  à travers le calcul du poids des  $N$  particules. Ainsi, si l'association de données est définie, l'étape suivante consiste à calculer les poids de toutes les particules qui appartiennent au

centre de gravité donné par (4.51). Alors, la liste de poids  $\mathbf{L}_k^{i,j}$  est calculée à partir de la distribution issue du mélange d'une loi gaussienne et d'une loi uniforme (cf. la figure 4.15) afin de garder toute l'information utilisée lors de l'association de données.  $\mathbf{L}_k^{i,j}$  est définie par :

$$\mathbf{L}_k^{i,j} = \{\psi(\mathbf{z}_k^j, \mathbf{x}_k^{i,l})\} \quad (4.53)$$

Finalement, les poids sont normalisés avant l'étape de rééchantillonnage.

### Calcul de la note de fusion

Dans le but d'estimer la confiance attribuée au résultat de la fusion de données déterminée par l'algorithme, chaque piste est associée à une note appelée *FCFD*, calculée à partir de trois critères :

- ▷ l'Age de la Piste (*AP*),
- ▷ la Redondance des Capteurs (*RC*),
- ▷ la Note de Détection Globale (*NDG*).

L'âge de la piste permet de déterminer si la cible détectée est suivie depuis longtemps ou non par le système. La redondance des deux capteurs permet d'évaluer si la cible détectée est bien le résultat ou non, d'une donnée fusionnée entre le capteur laser et le capteur vidéo. La note de détection globale fournie par les algorithmes mono-capteur permet de connaître la confiance attribuée à la classification de la cible détectée comme un piéton.

Les notes *AP* et *RC* sont calculées à partir d'une distribution gaussienne :

$$AP(t) = \begin{cases} \frac{1}{\sigma_o \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu_o}{\sigma_o}\right)^2\right] & 0 < t \leq \mu_o \\ 1 & t > \mu_o \end{cases} \quad (4.54)$$

où  $t$  est l'âge de la piste.  $\mu_o$  représente le temps minimum de maintenance en vie d'une piste lorsqu'elle n'est associée à aucune observation.  $\sigma_o$  permet de diminuer plus ou moins vite la valeur *AP*.

$$RC(x) = \frac{1}{\sigma_f \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_f}{\sigma_f}\right)^2\right] \quad (4.55)$$

Avec  $\mu_f$  le taux théorique de détection d'une cible par le capteur laser et le capteur vidéo durant une seconde.  $\sigma_f$  permet de diminuer ou d'augmenter plus ou moins vite la valeur *RC*. Enfin,  $x$  est le taux de détection de la cible suivie envoyé par le capteur laser et le capteur vidéo.

Finalement, la note *FCFD* est donnée par :

$$\alpha_k^j = \frac{AP + RC + \lambda_k^j}{3} \quad (4.56)$$

L'algorithme 10 explicite les différentes étapes de calcul.

## 4.3 Conclusion

Nous avons présenté dans la première partie de ce chapitre un algorithme de détection et de suivi de piéton fondé sur l'utilisation d'un unique capteur quatre plans laser. A cette fin, nous avons proposé une approche originale consistant à extraire des « objets-piétons »

---

**Algorithme 10** Fusion centralisée non paramétrique des données.
 

---

1. Pour  $k = 0$ , générer  $N$  échantillons à partir de chaque mesure  $j = 1, \dots, m_k$ , i.e  $\{\mathbf{x}_0^{j,n}\}_{n=1}^N = \{\mathbf{x}_0^{1,n}, \dots, \mathbf{x}_0^{m_k,n}\}_{n=1}^N$  où  $\mathbf{x}_0^{j,n} = p(\mathbf{X}_0^j)$ .
  2. Calculer la matrice  $\mathbf{A}_k$  pour toutes les observations  $m_k$  et tous les objets fusionnés  $O$   
 si  $(d_{ij} \geq C)$   
 $\mathbf{A}_k = d_{ij}$  avec  $1 \leq i \leq O$  et  $1 \leq j \leq m_k$  où  $d_{ij}$  est la probabilité d'associations entre l'objet fusionné  $i$  et l'observation  $j$  en utilisant les  $N$  particules et la note  $NDG$  suivant l'équation (4.51).  
 sinon  
 $\mathbf{A}_k^{i,j} = 0$  alors  $\{\mathbf{x}_k^{i,n}\} = \{\mathbf{x}_{k-1}^{i,n}\}$  aller à l'item 5.
  3. Calcul des poids  $\pi_k^i = p(\mathbf{z}_k^j | \tilde{\mathbf{x}}_k^i)$  puis normalisation, i.e,  $\pi_k^n = \frac{\pi_k^n}{\sum_{l=1}^N \pi_k^l}$ .
  4. Générer un nouveau nuage de particules  $\{\mathbf{x}_k^{i,n*}\}_{n=1}^N$  après rééchantillonnage de  $\{\mathbf{x}_k^{i,n}\}_{n=1}^N$ , selon la probabilité  $Pr(\mathbf{x}_k^{i,n*} = \mathbf{x}_k^{i,n}) = \pi_k^l$ .
  5. Prédire les nouvelles particules,  $\mathbf{x}_{k+1}^{i,n} = f(\mathbf{x}_k^{i,n*}, \mathbf{v}_k)$ ,  $n = 1, \dots, N$  en bruitant chaque particule à partir du modèle d'évolution.
  6. Calculer pour chaque estimation  $i$  sa note de confiance  $FCFD$  (4.56)  $\alpha_k^j$ .
  7.  $k = k + 1$  aller à l'item 2.
- 

à partir du calcul de la fonction de vraisemblance fondé sur les méthodes d'estimation non paramétriques à noyau.

L'extraction d'« objets-piétons » repose sur l'analyse d'une coupe 2D de l'environnement observé. Dans le cas général de notre application, notamment en environnement urbain complexe, la grande variabilité en hauteur des objets rencontrés provoque des variations non négligeables des surfaces apparentes des objets présents dans une image de profondeur en fonction de sa hauteur d'enregistrement. Le calcul de la vraisemblance fondée sur la combinaison des « objets-piétons » issue des quatre nappes permet de surmonter cette difficulté. Cette méthode repose sur une connaissance théorique du nombre de nappes et d'impacts laser qu'un piéton doit renvoyer en fonction de sa hauteur, de sa largeur et de sa distance avec le capteur.

Par la suite une méthode de suivi des piétons a été exposée. En effet, s'il est absolument nécessaire de détecter les piétons présents sur la chaussée afin de les éviter, il est également indispensable de connaître leur comportement cinématique pour effectuer correctement leur évitement. Le suivi d'objets est un moyen d'améliorer la performance du suivi. Un algorithme de suivi fondé sur un « moteur de suivi » de type SIR couplé à un modèle d'évolution à mouvement circulaire est proposé et détaillé.

Enfin, dans la deuxième partie de ce chapitre, une extension de l'algorithme de suivi multi-piétons mono-capteur à un algorithme multi-capteurs est exposée. En effet, la fusion des données reste un problème de combinaisons des mesures délivrées par les différents algorithmes de détection. Le choix de ne pas synchroniser les données les unes par rapport aux autres a permis de conserver l'« esprit » d'association de l'algorithme de suivi mono-capteur. Les spécifications du projet LOVE contraignaient les systèmes mono-capteur à associer à chaque mesure une note de détection globale. L'utilisation d'une distribution mélangée d'une loi gaussienne et d'une loi uniforme a permis d'exploiter à la fois des

#### 4.3. CONCLUSION

---

informations sur la position des objets et la confiance que l'algorithme accorde à cette détection. Une fusion centralisée des données permet d'illustrer l'approche proposée.

Nous consacrons le dernier chapitre à une étude sur la pertinence de ces différents algorithmes développés à partir de résultats obtenus sur des données simulées et réelles.



---

## Chapitre 5

# Expérimentations et résultats

Dans le chapitre 4, nous avons décrit deux algorithmes spécifiques au problème de détection et de suivi des piétons. Le premier concerne l'utilisation d'un capteur laser quatre plans. Le second propose de combiner un capteur laser avec une caméra vidéo. Ce chapitre aborde maintenant la validation de ces algorithmes pour du suivi multi-piétons dans un environnement urbain complexe.

Les deux algorithmes ont été testés à partir de données collectées lors de deux campagnes d'expérimentations. La première campagne d'expérimentation a été effectuée sur le campus universitaire d'Aubière à bord d'un VIPA <sup>1</sup> du LASMEA. La deuxième a été réalisée par Renault dans le centre d'une ville de la banlieue parisienne.

Permettre une évaluation des performances entre différents algorithmes de détection est un problème récurrent en lui même car il signifie l'utilisation d'une base de données de référence également appelée « vérité terrain ». Dans la première section de ce chapitre, nous proposons une approche permettant d'obtenir une vérité terrain. Deux scénarios choisis parmi ces deux campagnes d'expérimentations ont été définis comme « vérité terrain » sur des critères qualitatifs (type de zone traversée) et quantitatifs (nombre de piétons qui apparaissent et disparaissent de la scène). Dans la première partie de ce chapitre, les résultats permettant d'analyser les performances de notre algorithme de détection sont obtenus à partir de cette base. Nous avons mis en évidence que combiner les « objets-piétons » situés dans les quatre plans laser permet d'améliorer le taux de détection et de diminuer le nombre de fausses alarmes. La suite des expérimentations a montré que décorer complètement le module de détection et le module de filtrage n'était pas la meilleure solution. Dans le but d'améliorer la détection et le suivi de piéton à partir d'un capteur quatre plans laser, nous présentons dans ce chapitre une stratégie permettant au module de filtrage de se renseigner directement dans l'image laser lorsque le module de détection ne lui fournit pas les mesures attendues.

Le choix d'un capteur quatre plans laser se justifie par sa capacité à délivrer avec précision l'information de profondeur. Cependant, une image de profondeur avec une résolution limitée présente une faiblesse inhérente : une information peu discriminante rendant la tâche de reconnaissance difficile. Pour surmonter cette difficulté, une fusion centralisée des mesures provenant d'algorithmes laser et vidéo est évaluée à la fin de ce chapitre.

---

<sup>1</sup>Véhicule Individuel Public Autonome

## 5.1 Obtention d'une vérité terrain

### 5.1.1 Objectifs

Notre but ici est de fournir des scénarios de référence, c'est-à-dire de fournir des scénarios avec la vérité terrain associée permettant ainsi d'évaluer les différents algorithmes de détection et de pouvoir les comparer entre eux.

Bien que l'utilisation d'une vérité terrain soit un atout important lors de l'évaluation d'un algorithme de détection (de piéton ou autres), on observe malheureusement, un manque voire une inexistence de vérité terrain associée à des scénarios de piétons provenant de données laser. Cela est certainement dû à la difficulté de cette tâche. En effet, la vérité terrain doit permettre d'évaluer objectivement les performances d'un algorithme de détection de piéton.

### 5.1.2 Méthodes existantes

Cette vérité terrain devra donc être obtenue par une méthode autre que les algorithmes eux-mêmes qui seront évalués. A notre connaissance, aucune méthode automatique fiable n'existe pour obtenir une vérité terrain à partir de scénarios de piétons fondés sur des données laser. Nous avons alors recensé deux principales méthodes permettant d'obtenir une vérité terrain :

- ▷ par obtention manuelle,
- ▷ par images de synthèse (ou scénarios simulés).

#### Obtention manuelle

La vérité terrain peut s'obtenir manuellement. En effet, il est possible, si les images laser sont accompagnées d'images vidéo, de trouver visuellement les piétons à détecter dans les images de profondeur. L'obtention manuelle est très souvent la seule méthode qui permet d'obtenir une vérité terrain « à la carte » c'est-à-dire ne comprenant que des piétons qui doivent être détectés suivant certains critères. Par exemple, à un instant  $k$  donné, un piéton peut être visible dans le champ de vue vidéo mais pas dans celui du LiDAR ou vice-versa. Il faut donc tenir compte de ce critère lors de l'établissement de la vérité terrain afin de ne pas demander à l'algorithme de détecter quelque chose qui est invisible d'un point de vue capteur. Aussi, dans le cadre du projet LOVE, certains piétons ne sont pas à détecter comme par exemple ceux dont le buste n'est pas complètement visible.

#### Images de synthèse

La modélisation théorique du capteur laser peut être utilisée de manière directe pour générer des images de profondeur à partir d'une scène de synthèse. Le scénario étant défini à l'avance par l'utilisateur, la vérité terrain est donc naturellement connue. Ce principe a été développé au laboratoire LIVIC <sup>2</sup> à travers le projet SiVIC <sup>3</sup> (cf. la figure 5.1). Actuellement le projet E'Motive <sup>4</sup> propose de contribuer à l'amélioration de SiVIC afin de développer un simulateur de systèmes de détection permettant de substituer une « vérité

---

<sup>2</sup>Laboratoire sur les Interactions Véhicules-Infrastructure-Conducteur

<sup>3</sup>Simulateur Véhicule-Infrastructure-Capteur

<sup>4</sup>Environment Modeling for Perceptive Intelligent Vehicles

terrain simulée » à une « vérité terrain réelle » pour la mise au point et la validation d'algorithmes de détection utilisant des capteurs extéroceptifs (caméras, radars, laser, etc.).

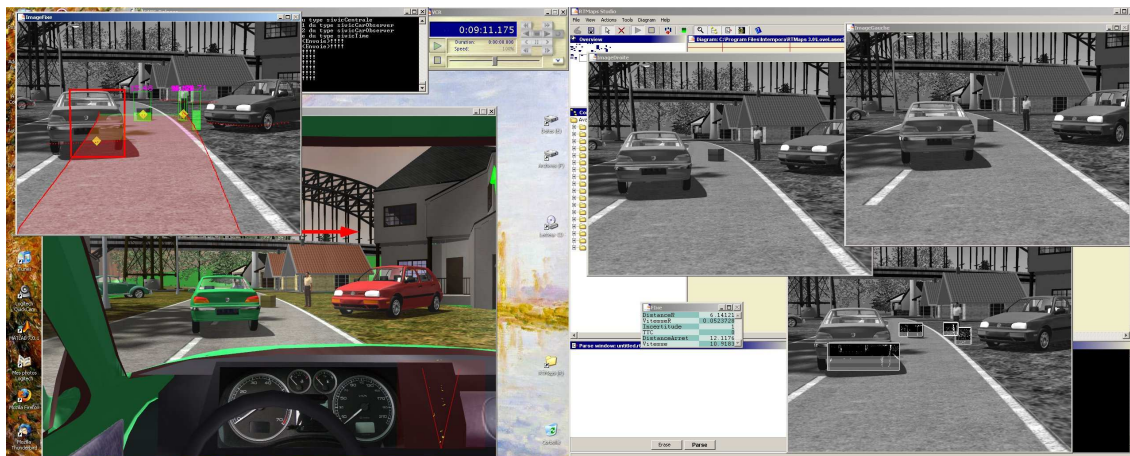


FIG. 5.1 – Exemples de scénarios produits par le simulateur SiVIC - figure extraite de [www.inrets.fr/ur/livic/](http://www.inrets.fr/ur/livic/).

### 5.1.3 Obtention manuelle de vérité terrain

Dans le cadre de nos évaluations, nous avons choisi d'utiliser la méthode d'obtention manuelle car même si cette méthode demande un travail long et fastidieux (il faut sélectionner à la main et dans chaque image, les piétons qui doivent être détectés), ce choix ne nécessite aucun nouveau développement de logiciel et permet à son utilisateur de créer une vérité terrain « à la carte ». Pour évaluer efficacement les algorithmes de détection de piéton proposés dans ce document, nous avons établi deux scénarios « vérité terrain » qui répondent à la question suivante : **Quels sont les piétons qui doivent obligatoirement être détectés ?**

Pour répondre à cette question, nous nous sommes appuyés sur le cahier des charges du projet LOVE que nous résumons ci-dessous :

- ▷ un piéton doit être absolument détecté si son buste est entièrement visible,
- ▷ tous les piétons en mouvement ou non doivent être détectés, sauf si la détection est en contradiction avec la première règle,
- ▷ tous les piétons orientés de face, de côté ou en arrière doivent être détectés, sauf si la détection est en contradiction avec la première règle,
- ▷ tous les piétons dont la taille est comprise entre un et deux mètres, doivent être détectés sauf si la détection est en contradiction avec la première règle,
- ▷ tous les piétons d'une largeur de buste comprise entre vingt et soixante centimètres doivent être détectés sauf si la détection est en contradiction avec la première règle ou la quatrième règle,
- ▷ tous les piétons doivent être détectés quelles que soient les conditions atmosphériques sauf si la détection est en contradiction avec la première règle, la quatrième ou la cinquième règle.

Les deux scénarios créés pour effectuer les évaluations de notre algorithme de détection fondé sur une combinaison d'images de profondeur, ne prennent donc pas en compte les piétons qui seraient invisibles dans l'image laser.



Nous présentons ci-dessous, les deux scénarios « vérité terrain » obtenus à partir de données enregistrées à bord du véhicule expérimental du LASMEA et celui de Renault.

- ▷ La figure 5.2 présente le scénario enregistré à bord d'un véhicule VIPA du LASMEA. Dans ce scénario, le véhicule évolue à une vitesse comprise entre 5 et 7  $km/h$  dans un environnement semi-urbain comportant notamment des arbustes, des bâtiments et quelques véhicules. Plusieurs piétons (jusqu'à trois simultanément) évoluent devant le véhicule qui suit une trajectoire rectiligne. De nombreuses occultations, apparitions et disparitions des piétons sont traités. Par souci de clarté dans la suite de ce chapitre, nous avons appelé ce scénario vérité terrain « **VIPA campus d'Aubière** ».

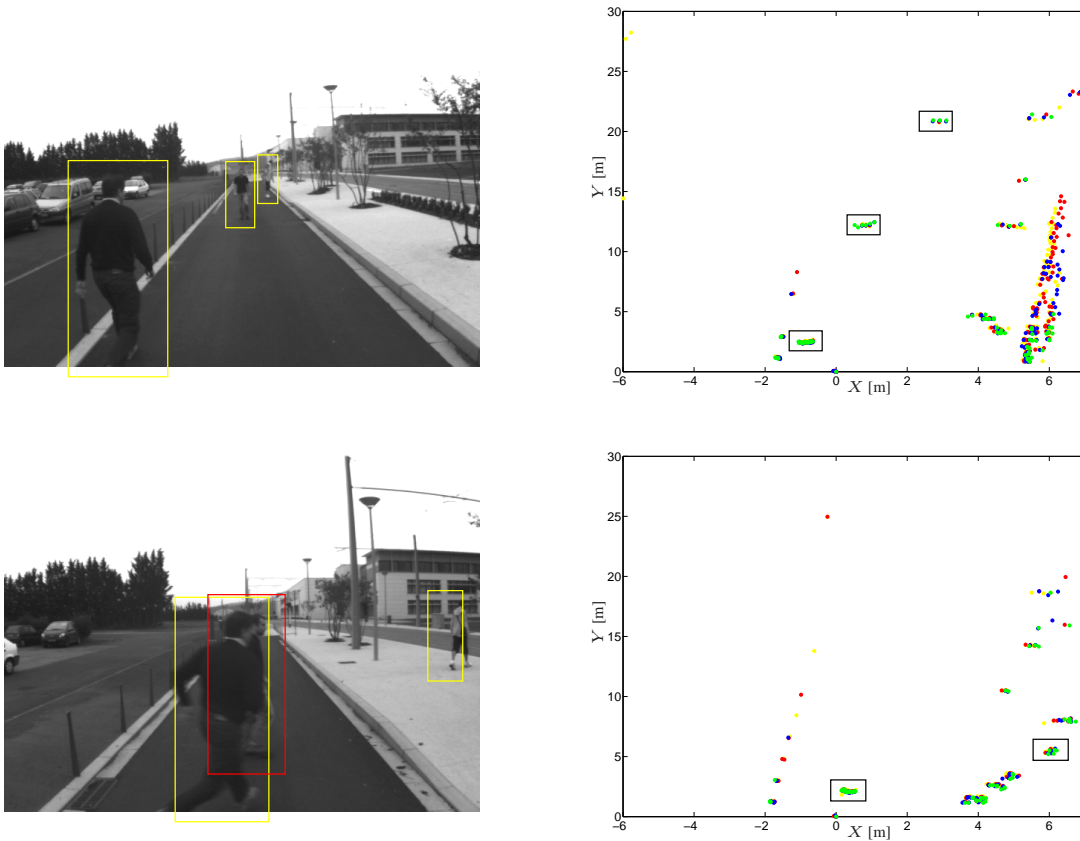


FIG. 5.2 – Extraits d'images vidéos et laser associées au scénario **VIPA campus d'Aubière**. Seuls les piétons encadrés en jaune font partie de la vérité terrain. À gauche : exemple d'une détection de trois piétons qui sont tous visibles au sens des spécifications LOVE. À droite : exemple d'un piéton encadré en rouge, visible ni dans la vidéo ni dans l'image de profondeur. Ce piéton n'appartient donc pas à la vérité terrain.

- ▷ La figure 5.3 présente le scénario enregistré à bord du véhicule expérimental de Renault. Dans ce scénario, le véhicule se déplace à une vitesse comprise entre 0 et 50  $km/h$  dans un environnement urbain comme ceux qui peuvent être rencontrés dans le centre-ville d'une grande agglomération française. Plusieurs piétons (jusqu'à cinq simultanément) évoluent dans la circulation comportant par exemple des véhicules et des camions. Les piétons sont à tout moment susceptibles d'être occultés par les autres usagers de la circulation. Par soucis également de clarté dans la

suite de ce chapitre, nous avons appelé ce scénario vérité terrain « **Renault centre-ville** » .

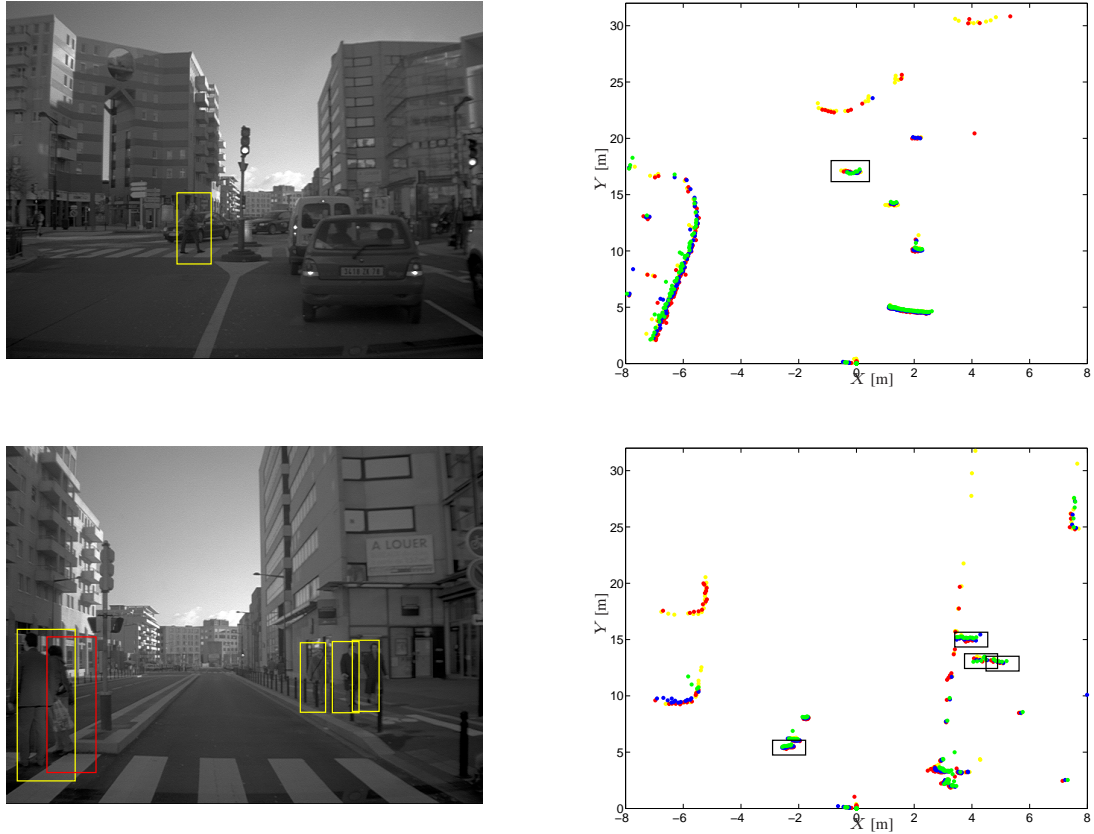


FIG. 5.3 – Extraits d’images vidéos et laser associées au scénario **Renault centre-ville**. Seuls les piétons encadrés en jaune font partie de la vérité terrain. A gauche : exemple d’une détection d’un piéton visible au sens des spécifications LOVE. A droite : exemple d’un piéton encadré en rouge, visible dans l’image de profondeur mais pas dans la vidéo. Ce piéton n’appartient donc pas à la vérité terrain.

### 5.1.4 Conclusion

Dans cette section nous avons présenté une méthode permettant de concevoir une vérité terrain à partir de critères précis donnés par le cahier des charges LOVE. Cette méthode propose d’établir manuellement et pour chaque image de profondeur, la liste des piétons à détecter. L’objectif de ce travail eonsiste à pouvoir mesurer objectivement les performances en détection de nos différents algorithmes par rapport à une même base de référence. Les deux scénarios choisis visent à évaluer la capacité des algorithmes à suivre un nombre variable d’« objets-piétons » dans différentes conditions de circulation, notamment vis-à-vis de l’environnement traversé (végétations, panneaux, barrières, etc.).

## 5.2 Comparaison de détection entre la méthode fondée sur la combinaison des quatre plans et la méthode fondée sur un seul plan

Plusieurs expérimentations ont été menées afin d'évaluer les performances de notre détecteur de piétons fondé sur l'utilisation d'un capteur quatre plans laser.

La première campagne d'expérimentation présentée dans cette section porte sur l'effet du recours à la combinaison des quatre plans, pour améliorer le taux de détection et diminuer le nombre de fausses alarmes. Nous avons mis en évidence dans le chapitre 4, la difficulté à assurer la détection des piétons à partir d'un unique plan de la scène observée car les apparences des objets sont souvent très proches.

En effet, une image de profondeur obtenue à une certaine hauteur ne fournit pas forcément les mêmes observations qu'une image prise à une autre hauteur car certains objets comme les plots d'intersections, les quilles et les bornes en ciment dépassent rarement les quatre vingts centimètres de haut.

De nombreux auteurs (cf. le chapitre 2) confrontés à cette difficulté, proposent de ne détecter que les piétons en mouvement, dans le but de filtrer les détections des nombreux objets statiques présents dans le champ de vision.

Pour éviter de filtrer les piétons immobiles, nous proposons dans le chapitre 4 de combiner les détections d'« objets-piétons » issues des quatre plans avant de prendre une décision finale sur la classification ou non d'un objet comme piéton. Cette méthode permet d'obtenir une information sur la hauteur des objets permettant ainsi de filtrer les petits objets de la circulation. Les performances respectives de la solution fondée sur une seule nappe avec celle fondée sur les quatre nappes sont comparées dans cette section vis-à-vis du nombre de bonnes et de fausses détections.

Ces expérimentations sont menées sur les deux scénarios vérité terrain présentés dans la section précédente. D'autres résultats sont ensuite présentés dans cette section afin d'illustrer la « généralité » de notre algorithme devant la diversité des environnements rencontrés.

### 5.2.1 Critères de performance de détection de piéton

Lors de l'étape précédente, une vérité terrain a été créée pour deux scénarios. Elle est représentée par un rectangle englobant les pieds du piéton.

L'évaluation consiste à mesurer les taux de bonnes détections et de fausses alarmes des algorithmes. Pour cela, on considère qu'une détection correspond à une vérité terrain lorsque la distance  $d$  est inférieure à la largeur de la détection (cf. la figure 5.4).

Les performances de détection sont alors évaluées grâce à deux critères :

- ▷ le taux de bonnes détections de piétons :  $\theta_B = \frac{N_D}{N_{VT}}$
- ▷ le taux de fausses alarmes :  $\theta_F = \frac{N_F}{N_T}$

avec  $N_T$  le nombre total de détections,  $N_D$  le nombre de piétons détectés,  $N_F$  le nombre de fausses détections et  $N_{VT}$  le nombre de piétons qui sont effectivement dans le champ du capteur (obtenu manuellement lors de l'établissement de la vérité terrain).

Les taux de bonnes et de fausses détections sont calculés pour chaque image de profondeur. La moyenne et l'écart-type par rapport à la vérité terrain du scénario complet sont obtenus par :

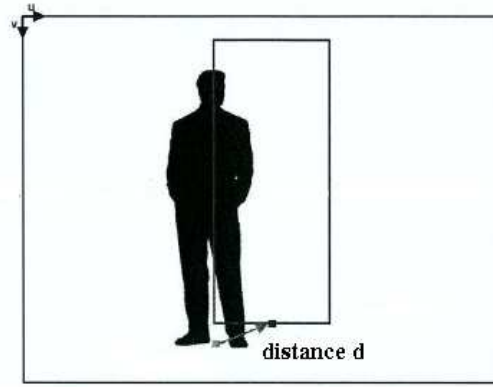


FIG. 5.4 – Illustration de la vérité terrain.

$$\bar{\theta} = \frac{1}{K} \sum_{k=1}^K \theta_k \quad (5.1)$$

$$\sigma_{\theta} = \sqrt{\frac{1}{K} \sum_{k=1}^K (\theta_k - \bar{\theta})^2} \quad (5.2)$$

### Résultats sur les deux scénarios vérité terrain VIPA campus d'Aubière et Renault centre-ville

Le code est écrit en langage C++ et utilise la bibliothèque NT<sup>2</sup> C++<sup>5</sup>. Nous utilisons un PC avec processeur Intel T2300 Core 2 Duo cadencé à 1,66 GHz avec 1Go de RAM, sous windows. Les traitements successifs des quatre nappes laser sont calculés en parallèle, par des threads fournis par la bibliothèque Boost C++<sup>6</sup>. Le traitement d'une image laser nécessite sur ce type de matériel d'un temps d'exécution entre 25 et 100 ms. En effet, ce temps est fonction du nombre d'impacts laser à traiter.

Toutes les expérimentations menées à bord du véhicule de Renault ont été effectuées en conditions réelles de circulation, c'est-à-dire dans le centre d'une ville de la banlieue parisienne avec un capteur laser IBEO Alasca réglé à une résolution angulaire de 0.25°.

Toutes les expérimentations effectuées à bord d'un des véhicules VIPA du LASMEA ont été menées pour des raisons de sécurité à une vitesse inférieure à 7 km/h dans un environnement considéré comme semi-urbain. Afin de tester la robustesse de notre algorithme à des capteurs moins performants, nous avons réglé la résolution angulaire de l'IBEO LDML à 0.5°.

Tous les objets cerclés de noir dans l'image laser ou encadrés dans l'image vidéo sont le résultat de la détection de l'algorithme fondé sur la combinaison des quatre plans laser. Les détections résultantes de chaque nappe sont renseignées dans l'image vidéo par un point de couleur : le vert pour la nappe 1, le bleu pour la nappe 2, le jaune pour la nappe 3 et le rouge pour la nappe 4. Dans l'image vidéo, les rectangles verts sont les détections confirmées par la vérité terrain tandis que les rectangles rouges sont les fausses

---

<sup>5</sup>Numerical Template Toolbox. <http://nt2.sourceforge.net>

<sup>6</sup><http://www.boost.org>

détections. Pour ces expérimentations, la nappe de référence utilisée pour comparer les deux méthodes est la bleue, c'est-à-dire la nappe parallèle au sol.

La moyenne et la variance des taux de bonnes et de fausses détections par rapport à la vérité terrain sont indiquées dans les tableaux 5.1 et 5.2.

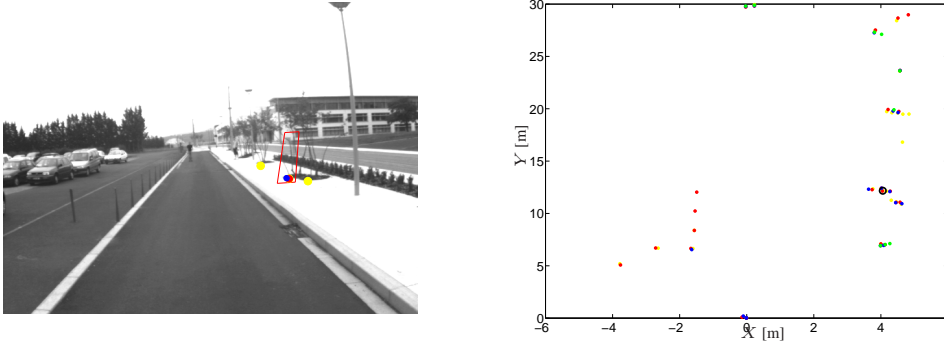


FIG. 5.5 – Relevés vidéo et laser pour l'image #1 du scénario **VIPA campus d'Aubière**. Ici, une fausse détection.

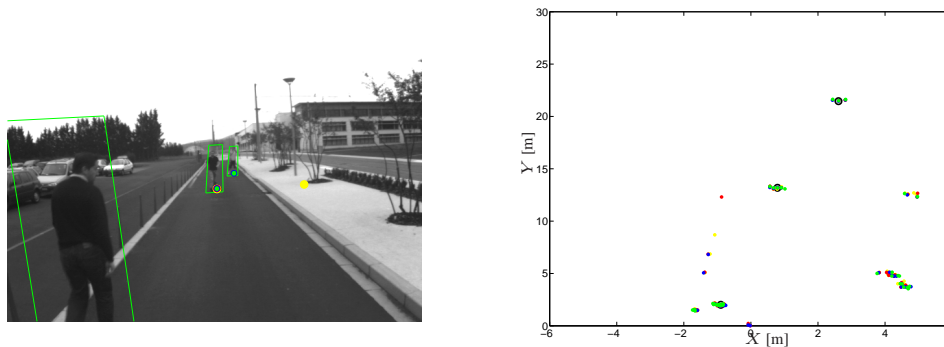


FIG. 5.6 – Relevés vidéo et laser pour l'image #160 du scénario **VIPA campus d'Aubière**. Ici, trois détections de piéton confirmées par la vérité terrain et aucune fausse détection.

TAB. 5.1 – Taux de fausses détections et de bonnes détections en fonction du nombre de nappes utilisées sur le scénario **VIPA campus d'Aubière**.

	un plan		4 plans	
	taux de bonnes détections	taux de fausses détections	taux de bonnes détections	taux de fausses détections
$\theta$	0.8124	0.188	0.9261	0.1749
$\sigma_\theta$	0.3382	0.3394	0.2156	0.3269

Le taux de fausses détections diminue beaucoup plus dans le scénario **Renault centre-ville** (- 16%) que dans le scénario **VIPA campus d'Aubière** (- 1%). La complexité nettement plus élevée du scénario **Renault centre-ville** comportant de nombreux obstacles

## 5.2. COMPARAISON DE DÉTECTION ENTRE LA MÉTHODE FONDÉE SUR LA COMBINAISON DES QUATRE PLANS ET LA MÉTHODE FONDÉE SUR UN SEUL PLAN

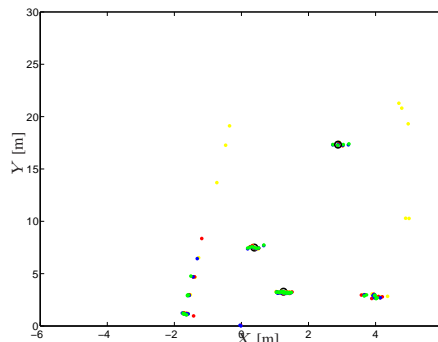


FIG. 5.7 – Relevés vidéo et laser pour l'image #179 du scénario **VIPA campus d'Aubière**. Ici, trois détections de piéton confirmées par la vérité terrain et aucune fausse détection.

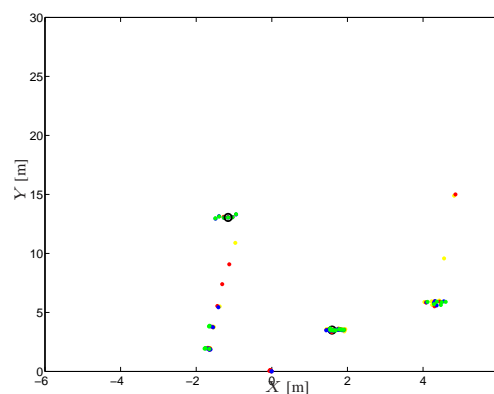


FIG. 5.8 – Relevés vidéo et laser pour l'image #218 du scénario **VIPA campus d'Aubière**. Ici, deux détections de piéton confirmées par la vérité terrain.

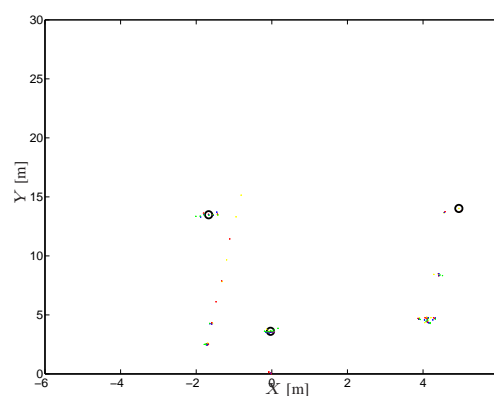


FIG. 5.9 – Relevés vidéo et laser pour l'image #228 du scénario **VIPA campus d'Aubière**. Ici, deux détections de piéton confirmées par la vérité terrain et une fausse détection.

rouitiers tels que des véhicules, panneaux, barrières... explique principalement cette diffé-



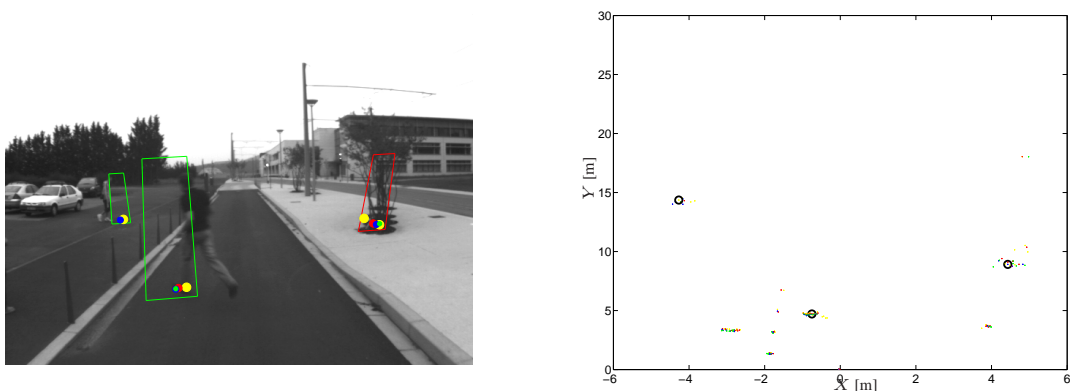


FIG. 5.10 – Relevés vidéo et laser pour l'image #268 du scénario **VIPA campus d'Aubière**. Ici, deux détections de piéton confirmées par la vérité terrain et une fausse détection.

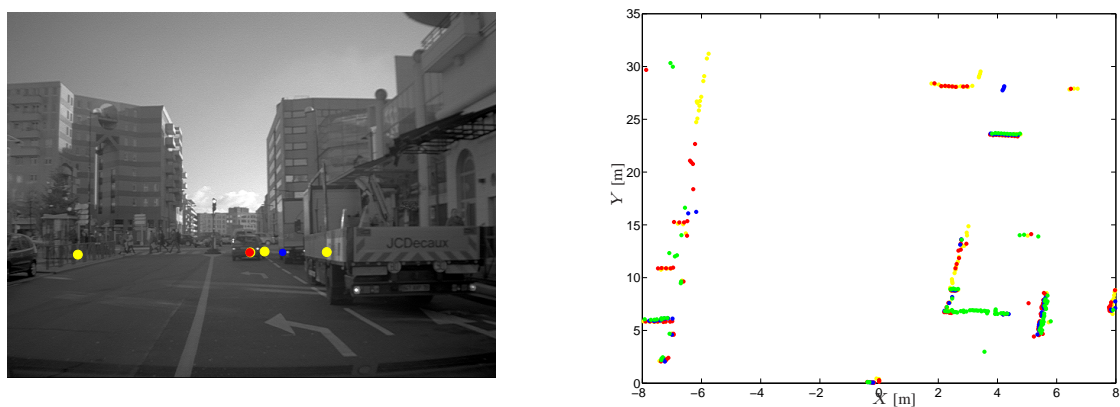


FIG. 5.11 – Relevés vidéo et laser pour l'image #1 du scénario **Renault centre-ville**. Ici, aucune détection de piéton et aucune fausse détection.

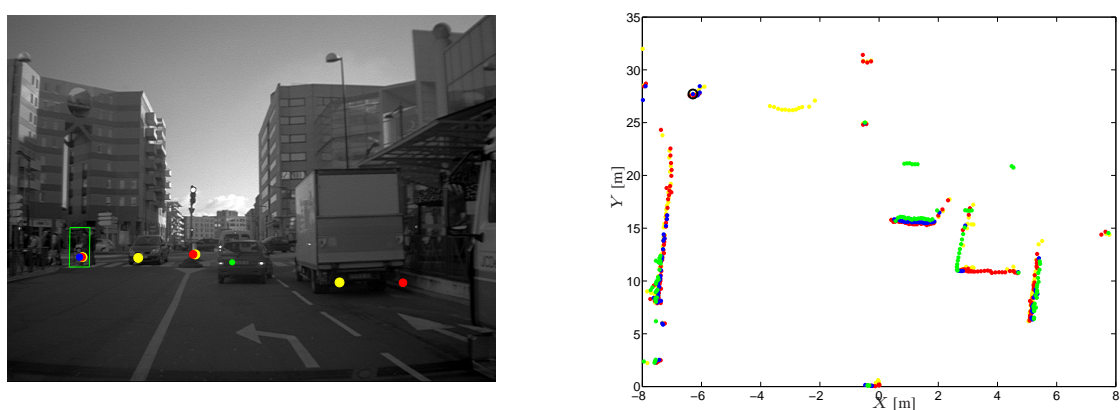


FIG. 5.12 – Relevés vidéo et laser pour l'image #11 du scénario **Renault centre-ville**. Ici, une détection de piéton confirmée par la vérité terrain.

## 5.2. COMPARAISON DE DÉTECTION ENTRE LA MÉTHODE FONDÉE SUR LA COMBINAISON DES QUATRE PLANS ET LA MÉTHODE FONDÉE SUR UN SEUL PLAN

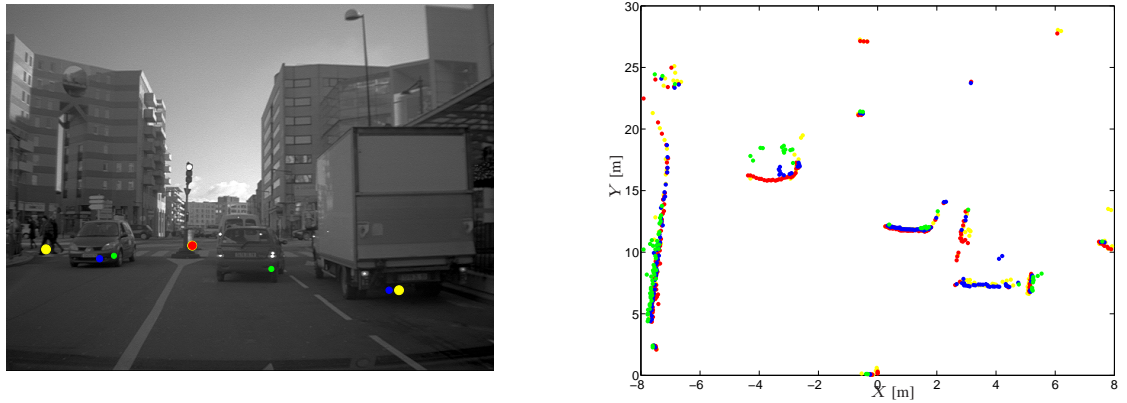


FIG. 5.13 – Relevés vidéo et laser pour l'image #16 du scénario **Renault centre-ville**. Ici, aucune détection de piéton et aucune fausse détection.

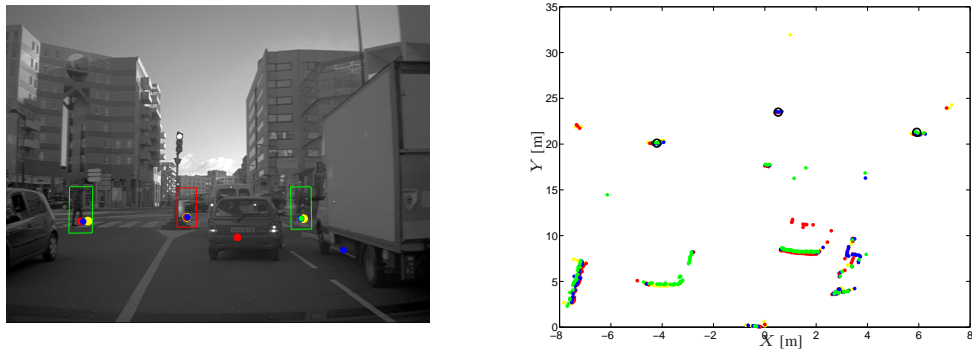


FIG. 5.14 – Relevés vidéo et laser pour l'image #22 du scénario **Renault centre-ville**. Ici, deux détections de piéton et une fausse détection.

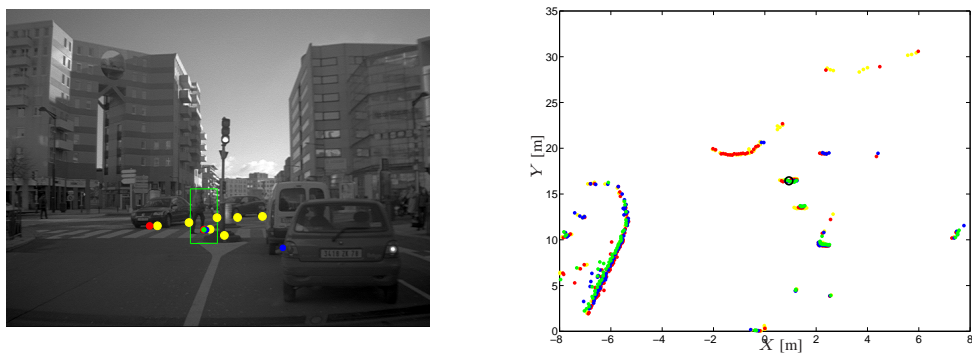


FIG. 5.15 – Relevés vidéo et laser pour l'image #32 du scénario **Renault centre-ville**. Ici, une détection de piéton et aucune fausse détection.

rence. En effet, l'algorithme mono-nappe est nettement plus sensible aux fausses détections provenant de ces nombreux objets que l'algorithme multi-nappes qui peut chercher confirmation. En revanche, le taux de bonnes détections s'améliore dans les mêmes proportions (+ 11%) sur les deux scénarios. Les causes sont :



TAB. 5.2 – Taux de fausses détections et de bonnes détections en fonction du nombre de nappes utilisées sur le scénario **Renault centre-ville**.

	un plan		4 plans	
	taux de bonnes détections	taux de fausses détections	taux de bonnes détections	taux de fausses détections
$\theta$	0.6163	0.6258	0.7289	0.4672
$\sigma_\theta$	0.3888	0.3894	0.2737	0.39

- ▷ des critères de détections identiques dans les deux vérités terrain,
- ▷ la complexité des scénarios n’influe pas sur l’amélioration du taux de bonnes détections entre un algorithme mono-plan et multi-plans car la brique d’extraction est la même dans les deux cas.

Seul le taux de bonnes détections final est différent (92% pour le scénario **VIPA campus d’Aubière** et 72% pour le scénario **Renault centre-ville**), car il est fonction de la complexité du scénario traité.

Les résultats obtenus à partir de la combinaison des quatre plans permettent de se rapprocher des besoins en détection des obstacles « piétons ». En revanche, la variance des taux de bonnes et fausses détections est très élevée (en moyenne 0.3). Deux explications simples peuvent l’expliquer :

- ▷ le nombre de piétons situés dans le champ de vue du capteur varie généralement entre 0 et 4. Le taux de bonnes détections peut donc simplement prendre cinq valeurs : 0, 0.25, 0.5, 0.75, 1. Avec un tel « delta » entre deux valeurs, la variance est souvent éloignée de la moyenne expliquant sa valeur élevée.
- ▷ pour des raisons similaires, le nombre de fausses détections excède rarement quatre à chaque image laser. La variance du taux de fausses détections est donc également souvent éloignée de la moyenne.

### Autres résultats expérimentaux de l’algorithme fondé sur la combinaison des quatre plans

Les expérimentations présentées dans cette section ont été menées dans diverses situations plus ou moins urbaines avec de nombreuses occultations entre objets. Ces résultats supplémentaires ont pour objectif de montrer quantitativement la « généricité » de l’algorithme de détection en s’adaptant à tout type d’environnement, dans différentes conditions de circulation à différentes résolutions angulaires (0.25° et 0.5°).

#### 5.2.2 Résultats expérimentaux de l’algorithme de suivi de piéton

La fonction d’un filtre est de produire une information temporellement plus cohérente que les mesures qui l’alimentent. Cette fonctionnalité permet par exemple de maintenir le suivi d’un piéton lorsque le module de détection défaille pendant quelques secondes. Les expérimentations ont été menées afin d’évaluer les performances de notre filtre SIR/KDE. Notamment, elles visent à valider la capacité de notre filtre SIR/KDE à suivre et identifier

## 5.2. COMPARAISON DE DÉTECTION ENTRE LA MÉTHODE FONDÉE SUR LA COMBINAISON DES QUATRE PLANS ET LA MÉTHODE FONDÉE SUR UN SEUL PLAN

un nombre variable de piétons dans différentes situations. Les fonctionnalités suivantes sont évaluées :

- ▷ précision du suivi,
- ▷ reconstruction de la vitesse,
- ▷ identification des piétons.

Cette section est organisée ainsi : dans un premier temps, nous présentons un exemple de scénario simulé permettant de valider les deux premiers points. Puis dans un second temps, nous proposons d'évaluer la cohérence temporelle des identités à partir d'un scénario de données réelles enregistré dans le garage du LASMEA.

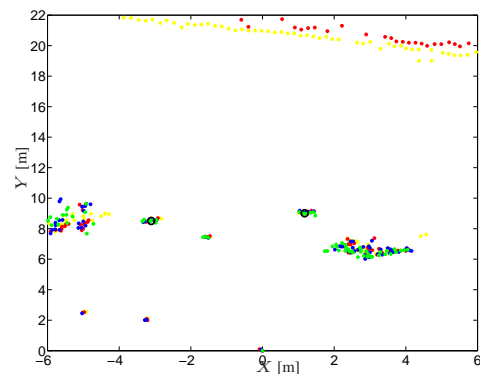
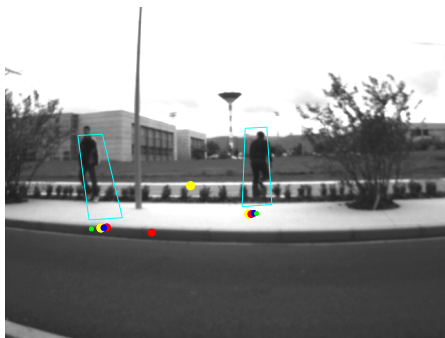
### Séquence simulée : évaluation de la précision en position du suivi et de la reconstruction de la vitesse

Les expérimentations de précision en position et vitesse visent à évaluer la capacité du filtre à suivre un piéton en position et vitesse. Pour ce faire, nous avons généré des séquences de synthèse spécifiques. Le scénario présenté ici, simule la trajectoire d'un piéton situé à environ six mètres du capteur laser en se déplaçant horizontalement de la gauche vers la droite, à une vitesse constante arbitrairement fixée à  $2\text{ m/s}$  (cf. la figure 5.18). Les données sont supposées non bruitées et envoyées toutes les  $150\text{ ms}$ . Lors de ces simulations, le filtre SIR utilisait 400 particules et un modèle d'évolution à mouvement circulaire. Dans le contexte du suivi des piétons,  $\sigma_p = 2\text{ m.s}^{-2}$  est l'écart-type des accélérations en profondeur,  $\sigma_\theta = 5^\circ.\text{s}^{-2}$  est l'écart-type des accélérations en angle et  $\sigma_g = 10^\circ$  est l'écart-type du bruit sur le cap de la trajectoire du piéton. Le scénario a été répété 10 fois. L'erreur en position et vitesse est illustrée dans la figure 5.20.

Les résultats illustrés par les figures 5.18, 5.19, 5.20 et 5.21 montrent que l'algorithme de suivi donne de bons résultats en position avec une erreur en profondeur toujours inférieure à 0.02 mètres et inférieure à 0.2 degrés en angle. L'algorithme démontre aussi une bonne capacité à estimer la vitesse autant en profondeur qu'en angle.

### Données réelles

Le scénario présente le suivi de deux piétons situés dans un environnement intérieur en utilisant l'algorithme de détection basé sur les quatre nappes laser. Lors de ces expérimentations, le filtre SIR utilisait 400 particules et un modèle d'évolution à mouvement circulaire. Dans le contexte du suivi des piétons,  $\sigma_p = 2\text{ m.s}^{-2}$  est l'écart-type des accélérations en profondeur,  $\sigma_\theta = 5^\circ.\text{s}^{-2}$  est l'écart-type des accélérations en angle et  $\sigma_g = 10^\circ$



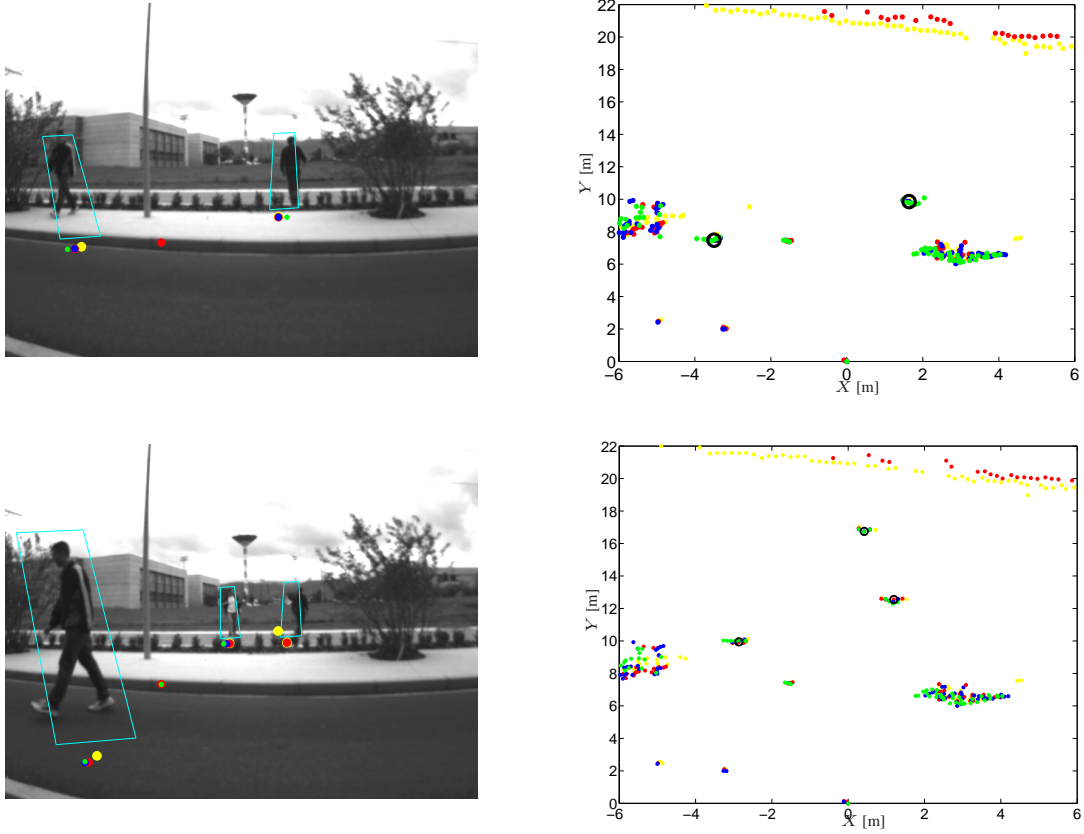
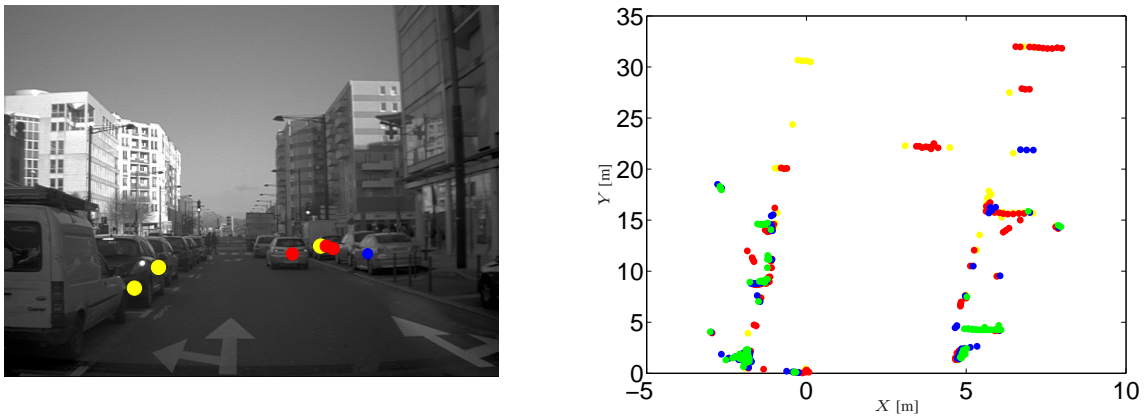


FIG. 5.16 – Le premier scénario présente le suivi de trois piétons situés dans un environnement semi-urbain en utilisant l’algorithme de détection fondé sur les quatre nappes laser. Ce scénario est enregistré à bord du véhicule expérimental du LASMEA. Les principales difficultés de ce scénario sont le filtrage des arbustes et les fréquentes occultations des piétons entre eux. Les piétons détectés dans l’image laser sont cerclés de noir. Les rectangles bleus représentent leur projection dans l’image vidéo.



est l’écart-type du bruit sur le cap de la trajectoire du piéton. Les piétons détectés dans l’image laser sont cerclés de noir. Les rectangles de couleur représentent leur projection dans l’image vidéo. Les différentes couleurs des rectangles représentent l’identité des

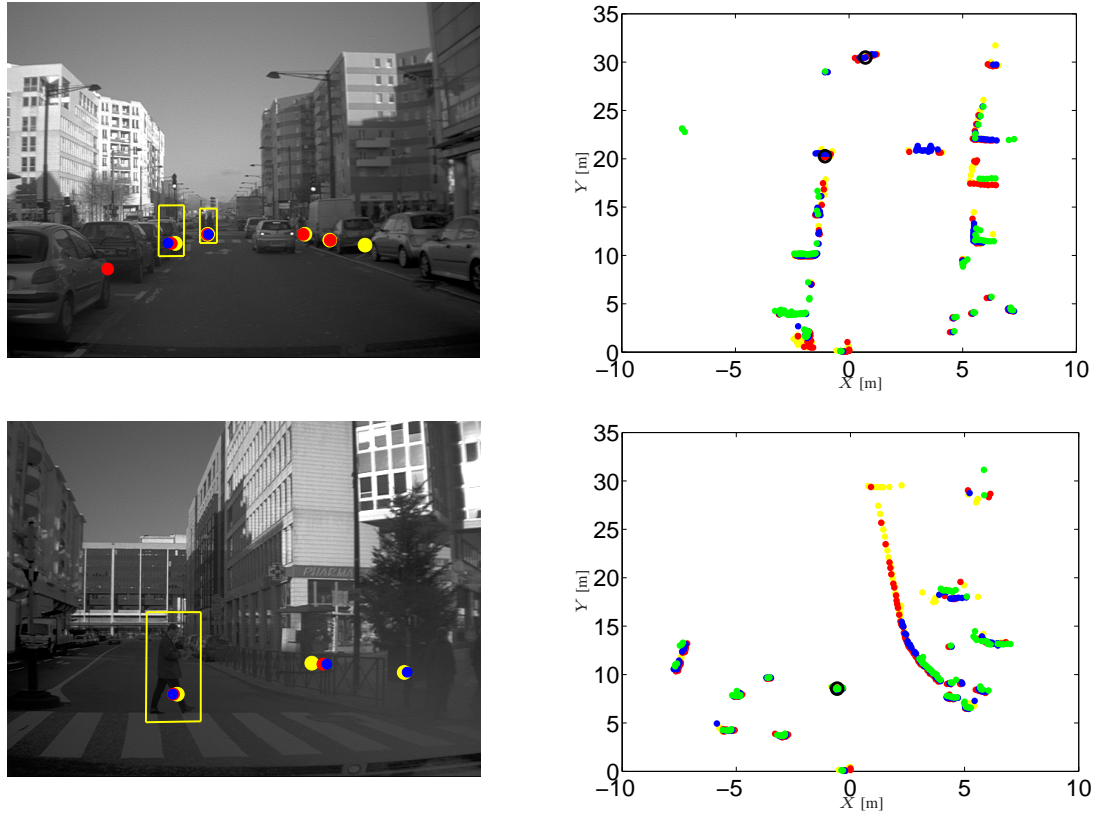


FIG. 5.17 – Le second scénario présente le suivi de plusieurs piétons situés dans un environnement urbain en utilisant l’algorithme de détection fondé sur les quatre nappes laser. Ce scénario a été enregistré à bord du véhicule expérimental de Renault dans des conditions de circulation à vitesse réelle. Les principales difficultés rencontrées sont le filtrage des gros poteaux de circulation, les véhicules garés sur les bas-côtés, les fréquentes occultations des piétons entre eux et de manière générale, la richesse de l’environnement urbain qui induit de nombreuses détections assimilables à un piéton. Nous reverrons ce scénario plus en détails lors de l’évaluation de l’algorithme de fusion centralisée laser/vidéo. Les piétons détectés dans l’image laser sont cerclés de noir. Les rectangles blancs représentent leur projection dans l’image vidéo. On rappelle que les points de couleur présents dans l’image vidéo représentent la projection des détections effectuées dans les quatre plans.

pistes « piétons » qui sont suivies.

### 5.2.3 Conclusion

Dans cette section, nous avons évalué dans un premier temps les performances de détection des piétons de l’algorithme proposé dans le chapitre 4. Ce travail a permis de démontrer que l’utilisation judicieuse des quatre plans laser permet de mieux détecter des piétons et de diminuer de manière significative le nombre de fausses détections. Nous considérons à ce stade de l’étude que la fusion des informations contenues dans les quatre plans laser, grâce à la méthode non paramétrique fondée sur les noyaux, permet de sélectionner efficacement l’amas d’impacts laser possédant les caractéristiques géométriques d’un piéton.

Dans un second temps, nous avons évalué les performances de pistage de notre filtre

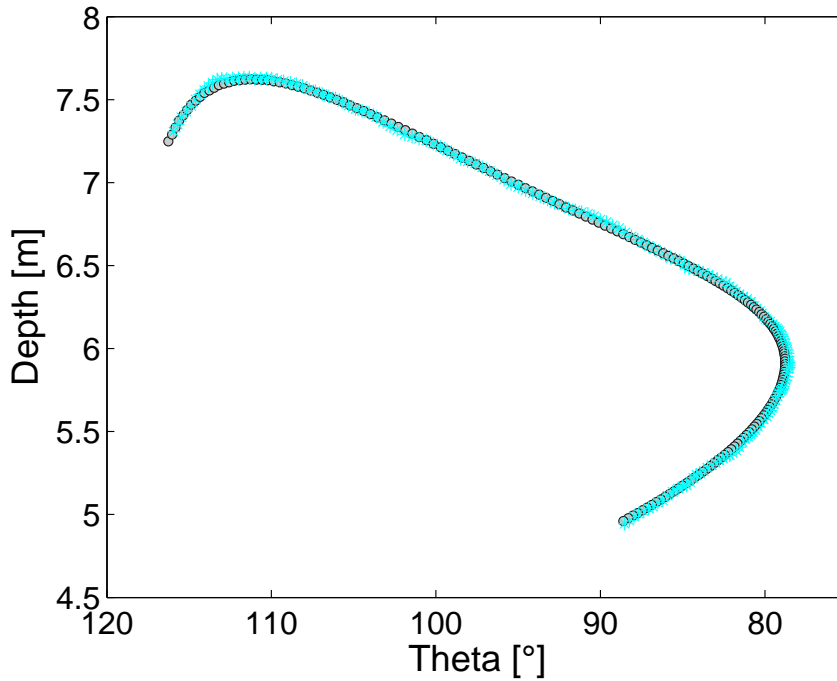


FIG. 5.18 – Résultat du suivi d'un piéton avec l'algorithme SIR/KDE : vue de dessus. Les mesures sont représentées par des cercles gris.

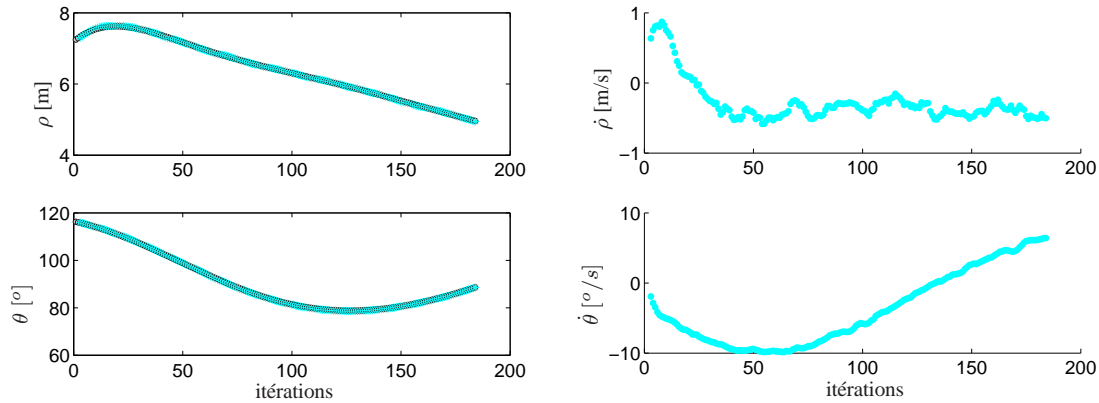


FIG. 5.19 – Résultats en positions ( $\rho$  et  $\theta$ ) du suivi d'un piéton. Les mesures sont représentées par des cercles gris.

SIR/KDE. Les précisions en position et vitesse sont satisfaisantes et répondent parfaitement aux besoins du projet. En revanche sa résistance aux périodes d'invisibilité n'est pas satisfaisante. En effet, entre l'image 35 et 113, le piéton situé au fond de la salle a été occulté pendant un certain temps ( $\sim 1.2s$ ) par le piéton au premier plan (image 47) devenant fortement voire totalement invisible dans l'image de profondeur. N'étant plus détecté depuis plus de  $0.5s$ , le module de suivi décide de détruire sa piste (identifiée rouge) et de générer une nouvelle piste (identifiée verte) à sa réapparition.

Il est donc nécessaire d'améliorer le point faible d'une association séquentielle module de détection puis module de suivi. La section suivante vise à apporter une réponse à ce problème.

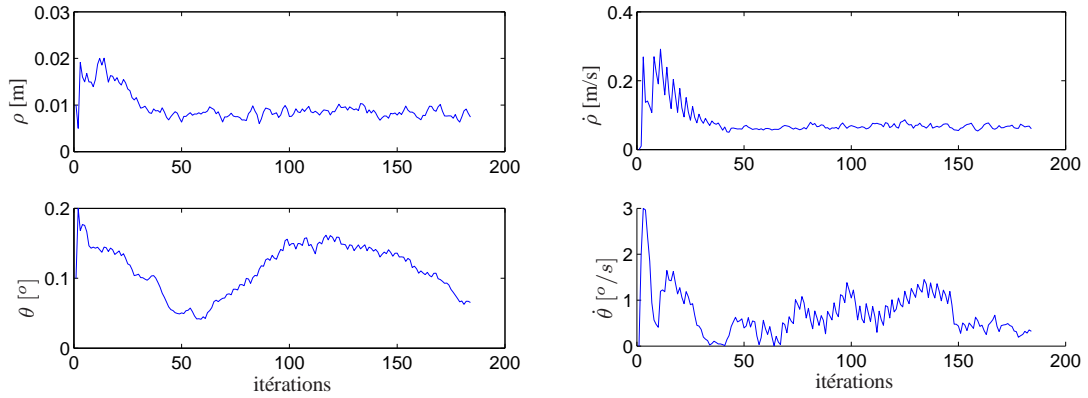
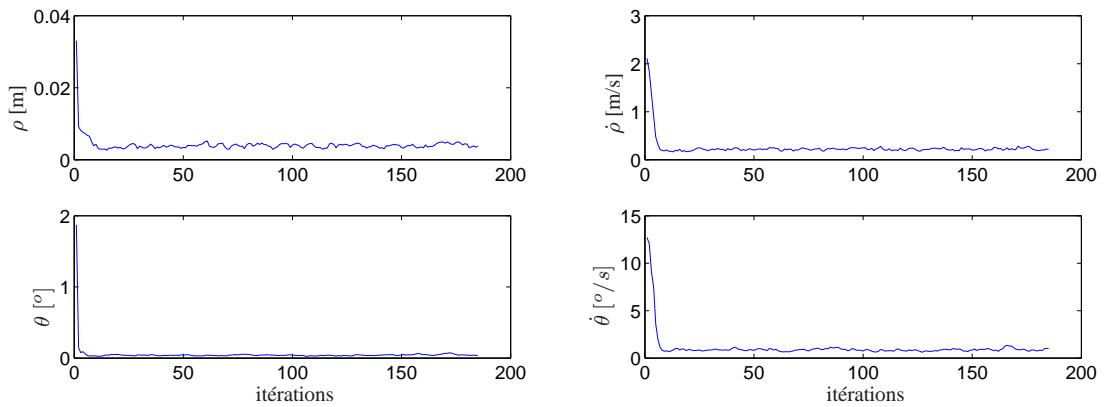
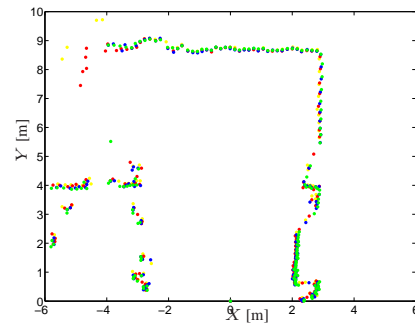

FIG. 5.20 – Résultats de l'erreur en positions ( $\rho$  et  $\theta$ ) du suivi d'un piéton.

FIG. 5.21 – Résultats de la covariance en positions ( $\rho$  et  $\theta$ ) du suivi d'un piéton.


FIG. 5.22 – Relevés vidéo et laser pour l'image #1.

## 5.3 Détection et suivi de piéton avec retour dans l'image laser

### 5.3.1 Problématique

Nous avons vu à la fin de la section précédente que lorsque le module de détection défaille dans sa mission à cause d'une observabilité provisoirement mauvaise (image mal

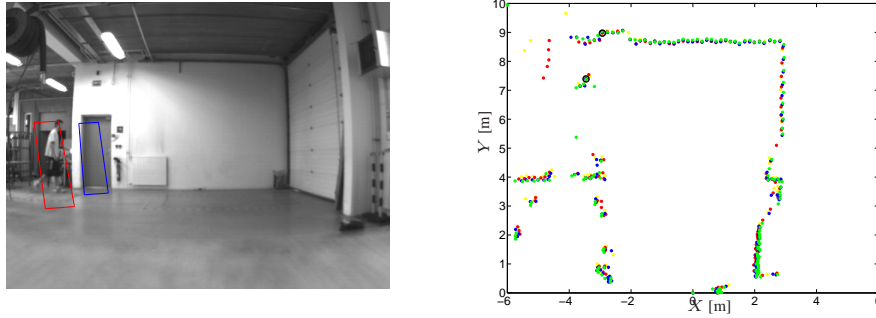


FIG. 5.23 – Relevés vidéo et laser pour l'image #5.

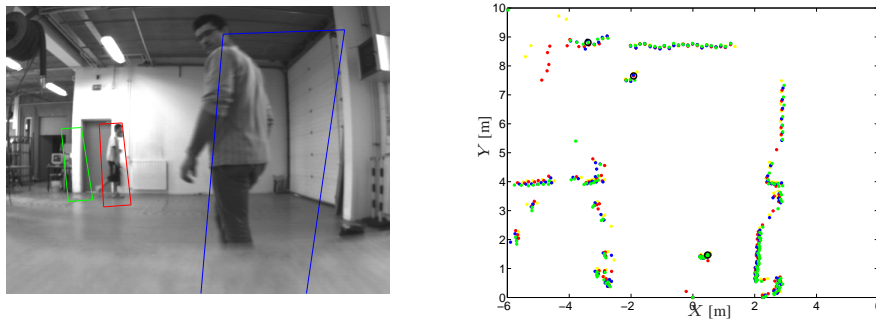


FIG. 5.24 – Relevés vidéo et laser pour l'image #35.

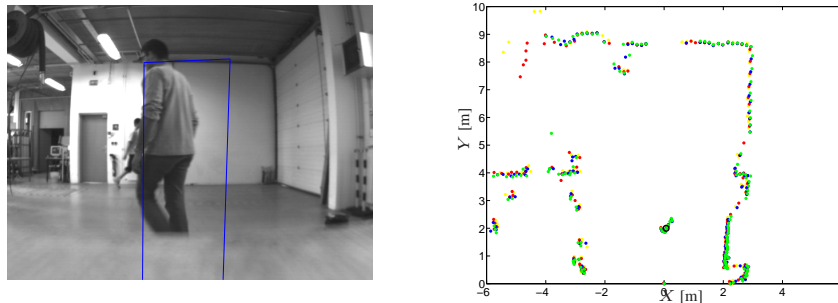


FIG. 5.25 – Relevés vidéo et laser pour l'image #47.

segmentée, objets occultés,...), le module de suivi situé en aval, se retrouve alors sous alimenté en mesures et propose de détruire des pistes qui correspondent à des objets toujours présents dans le champ de vision du capteur.

Nous proposons dans cette section, d'ajouter une boucle de retour (en anglais : *feedback*) entre le module de suivi et l'image laser permettant ainsi au filtre de vérifier par lui-même la non observabilité de l'objet suivi. En effet, le système actuel présente l'inconvénient de rendre complètement dépendant le module de suivi aux mesures fournies par le module de détection et donc à ses défaillances.

Cette boucle a donc pour but d'améliorer la cohérence temporelle du nombre d'objets



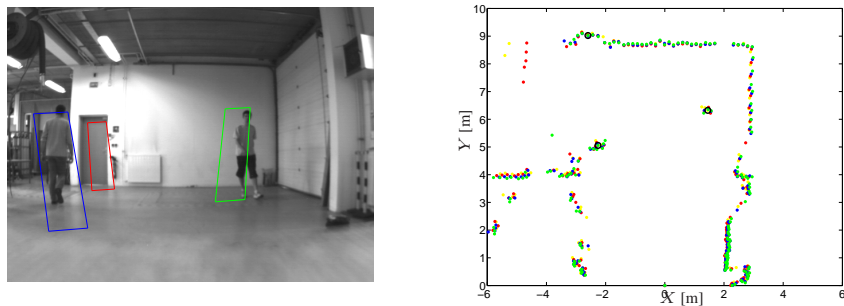


FIG. 5.26 – Relevés vidéo et laser pour l'image #113.

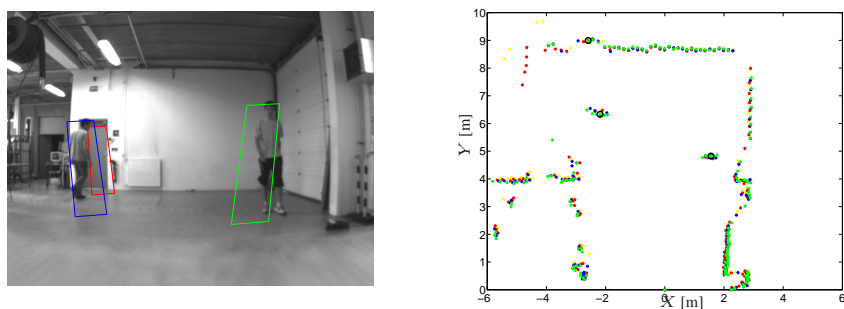


FIG. 5.27 – Relevés vidéo et laser pour l'image #139.

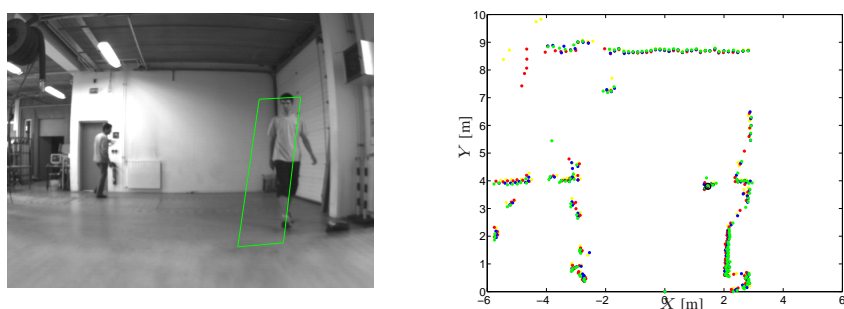


FIG. 5.28 – Relevés vidéo et laser pour l'image #157.

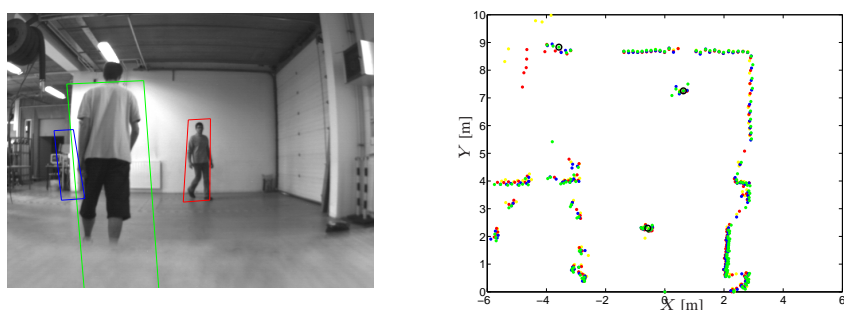


FIG. 5.29 – Relevés vidéo et laser pour l'image #234.



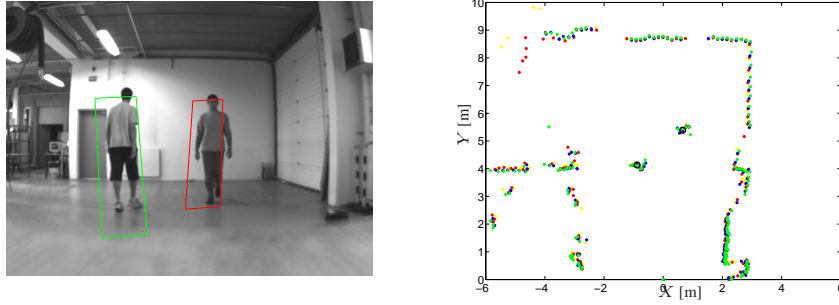


FIG. 5.30 – Relevés vidéo et laser pour l'image #278.

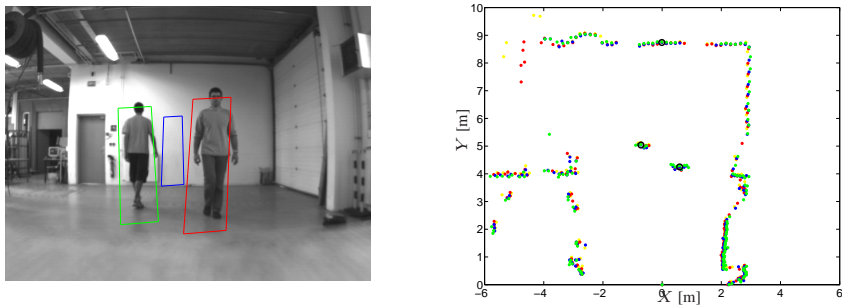


FIG. 5.31 – Relevés vidéo et laser pour l'image #296.

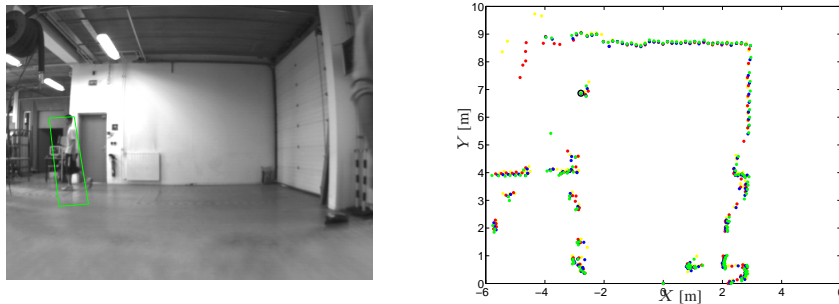


FIG. 5.32 – Relevés vidéo et laser pour l'image #345.

présents dans la scène observée, en vérifiant directement dans l'image laser si un amas de points ne pourrait pas s'associer avec une piste laissée sans mesure.

### 5.3.2 Du modèle classique au modèle avec rétroaction

Cette section présente l'algorithme de détection et de suivi de piéton avec *feedback* dans l'image laser (cf. la figure 5.33). Dans un premier temps, l'algorithme de segmentation utilisé pour construire un segment à partir d'un amas de points candidat est présenté. Dans un second temps, l'algorithme de filtrage avec *feedback* dans l'image de profondeur est détaillé. Le module de détection (en vert sur la figure 5.33) quant à lui n'est pas détaillé car identique à celui décrit dans le chapitre 4.

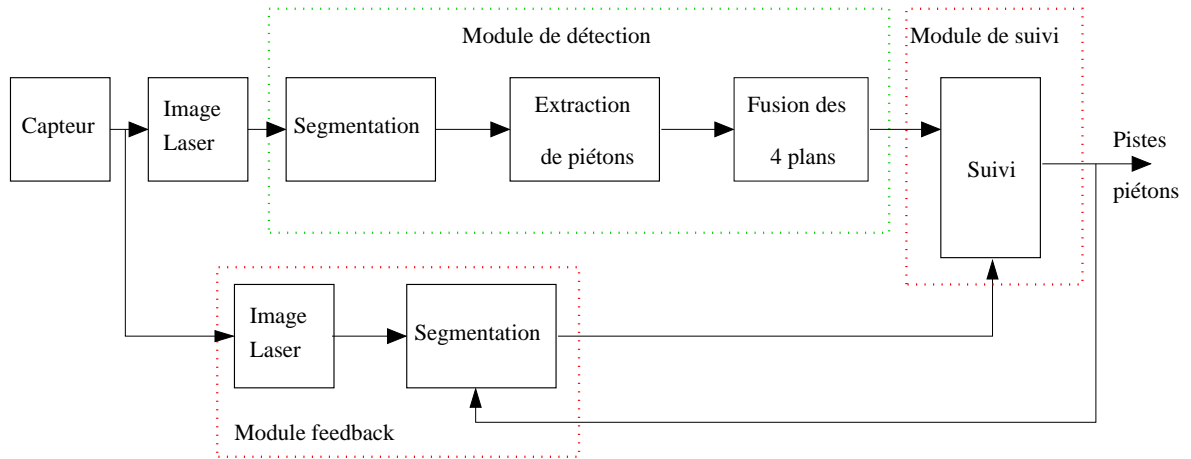


FIG. 5.33 – Schéma de l'algorithme de détection et de suivi de piéton avec *feedback* dans l'image laser.

### Module de suivi avec boucle de retour

Comme expliqué dans la section 5.3.1, nous allons utiliser la connaissance apportée par le module de filtrage pour aider la détection des piétons lorsque l'observabilité est mauvaise. Notre contribution est d'ajouter un lien direct entre le filtre et l'image de profondeur.

Lors de l'étape d'association de données, si un objet  $\mathbf{x}_k^i$  ne trouve pas de mesure  $\mathbf{z}_{k-1}^j$  dans l'ensemble des mesures  $\mathbf{Z}_{k-1}$ , le modèle de filtrage propose de vérifier directement dans l'image laser si un groupe de points pourrait expliquer l'objet  $\mathbf{x}_k^i$ . Dans le cas où un groupe de points  $\{\mathbf{M}_i = (x_i, y_i)\}_{i=1, \dots, N}$  existe à la position pointée par l'objet  $\mathbf{x}_k^i$ , nous utilisons l'algorithme de la corde [Ramer, 1972] afin de classifier le groupe de points comme appartenant ou non à la classe piéton. Cet algorithme est un algorithme classique et assez simple de segmentation. En effet, il modélise un objet par une suite de segments de manière itérative tant que l'erreur du modèle proposé est supérieure à un seuil à définir. Le nombre de segments est donc fonction du seuil de précision demandé sur le modèle.

Notre approche applique donc cet algorithme au groupe de points localisé par l'objet  $\mathbf{x}_k^i$ . Si ce groupe de points peut être décrit par un segment entrant dans le gabarit d'un piéton, ce segment est validé comme mesure. Cette mesure est alors intégrée dans le vecteur des mesures  $\mathbf{Z}_{k-1}$  qui est ensuite envoyé au module de filtrage qui lancera une nouvelle étape d'association de données.

### Résultats expérimentaux de l'algorithme de suivi avec *feedback*

L'ensemble de l'algorithme est écrit sous Matlab. Pour une comparaison objective, tous les résultats sont obtenus avec le même jeu de paramètres que l'algorithme sans *feedback*.

Afin de valider notre approche et de montrer sa pertinence, nous proposons d'utiliser une courbe ROC (Receiver Operating Characteristic) afin de permettre de comparer les performances des deux algorithmes de suivi (avec ou sans *feedback* dans l'image laser) (cf. la figure 5.35). En effet, la courbe ROC est un outil qui permet de comparer les algorithmes d'apprentissage ou de classification. Elle indique la capacité du classifieur à placer les positifs devant les négatifs en mettant en relation sur le graphique les taux de

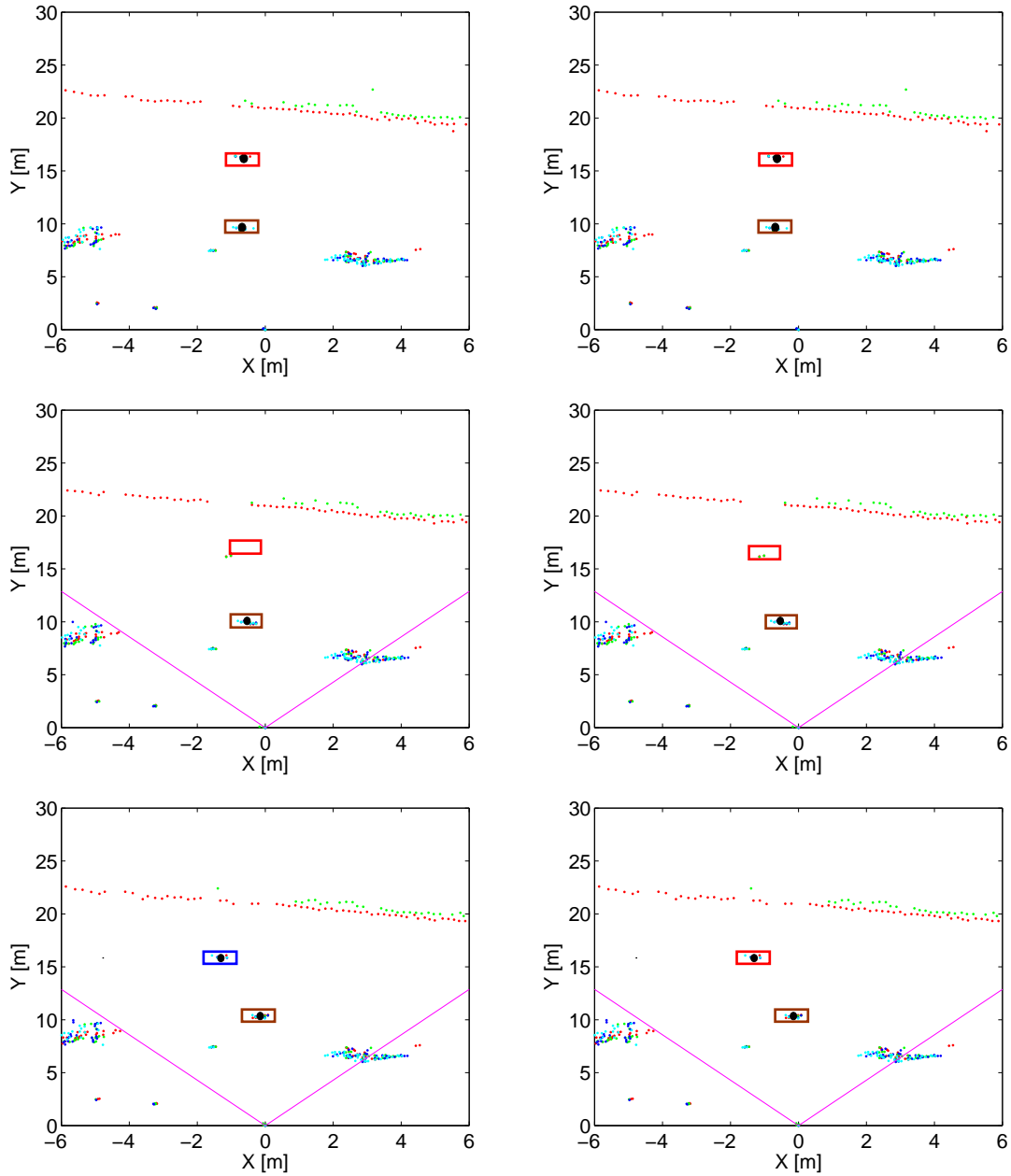


FIG. 5.34 – Les piétons détectés sont représentés par des points noirs tandis que les piétons suivis sont représentés par des rectangles de couleur. La couleur permet d'identifier la piste. A gauche, extrait d'un suivi de piéton dans une image laser avec l'algorithme sans boucle de retour. Ici, quand le piéton en second plan est occulté, le module de pistage sans *feedback* perd sa trace car il n'est pas détecté par le module de détection. Lorsque ce piéton redevient observable par le module de détection, le module de suivi propose d'initialiser une nouvelle piste, munie d'une nouvelle identité (rectangle qui passe du rouge au bleu). A droite, même extrait avec l'algorithme utilisant la boucle de retour. Cette fois, le module de pistage avec *feedback* arrive à garder une trace du piéton en le cherchant dans l'image laser permettant d'obtenir une mesure capable de s'associer avec la piste qui survit et conserve son identité (pas de changement de couleur du rectangle).

---

**Algorithme 11** Module de détection avec *feedback* dans l'image laser.

---

1. Calculer l'association de données entre tous les objets  $i$  et toutes les mesures  $j$   
**si** (tous les objets sont associés à une mesure)  
     Aller à l'item 4  
   **sinon**  
     Envoyer dans l'image laser les positions des objets  $i$ .
  2. Rechercher les amas de points susceptibles de s'associer avec les objets  $x_k^i$ .  
**si** (nombre de points laser  $\leq 2$ )  
     Pas de nouvelle mesure, aller à l'item 4  
   **sinon**  
     Algorithme de la corde.
  3. Classification des segments obtenus.  
**si** (taille  $>$  gabarit minimum)  
     Créer une nouvelle mesure  $z_k^{j+1}$   
   **sinon**  
     Pas de nouvelle mesure.
  4. Algorithme « Association et suivi temporel : SIR/KDE » présenté au chapitre 5.
  5.  $k = k + 1$  aller à l'item 1.
- 

faux positifs (en abscisses) et les taux de vrais positifs (en ordonnées).

#### Construction de la courbe ROC

Pour construire cette courbe, nous faisons varier la valeur du seuil de confiance  $\delta = [0, 1]$ . Pour chacune de ces variations, on calcule le taux de bonnes détections  $\bar{\theta}_B$  et le taux de fausses détections  $\bar{\theta}_F$  sur le scénario **VIPA campus D'Aubière**. Chaque valeur du seuil  $\delta$  est alors associée à un couple  $(\bar{\theta}_B, \bar{\theta}_F)$  que l'on reporte dans un graphique sur un plan orthonormé où les abscisses ( $x$ ) représentent le taux de fausses détections variant de 0 à 1 et où les ordonnées ( $y$ ) représentent le taux de bonnes détections.

#### Interprétation de la courbe ROC

La courbe ROC (cf. la figure 5.35) montre que le détecteur avec *feedback* possède un pouvoir discriminant plus fort (courbe rouge au-dessus de la verte) que le détecteur sans *feedback*. Notons que ce pouvoir discriminatoire s'accroît principalement dans la première partie de la courbe c'est-à-dire lorsque le seuil de confiance  $\delta$  tend vers 1. Dans cette configuration, le module de détection devient très sélectif permettant de diminuer le nombre de fausses détections mais également le nombre de bonnes détections. Cette dernière remarque vérifie le rôle de la boucle de retour qui permet de pallier aux défaillances d'observabilité du module de détection. En effet, à partir de 50% de faux positifs, le pouvoir discriminatoire devient sensiblement identique avec les deux algorithmes.

Les figures 5.36 et 5.37 sont très intéressantes car elles illustrent en vue de dessus les résultats du suivi de deux piétons en environnement intérieur, scénario commenté dans la section 5.2.2. Nous rappelons que le principal problème est le changement d'identité du piéton situé au second plan (au fond de la salle) lorsqu'il est occulté par le piéton situé au premier plan. La figure 5.37 montre l'avantage de l'algorithme de suivi avec *feedback* permettant de ne pas perdre le suivi du piéton situé au fond de la salle (piste verte) alors que l'algorithme sans *feedback* (cf. la figure 5.36) l'avait perdu (passage de la piste rouge

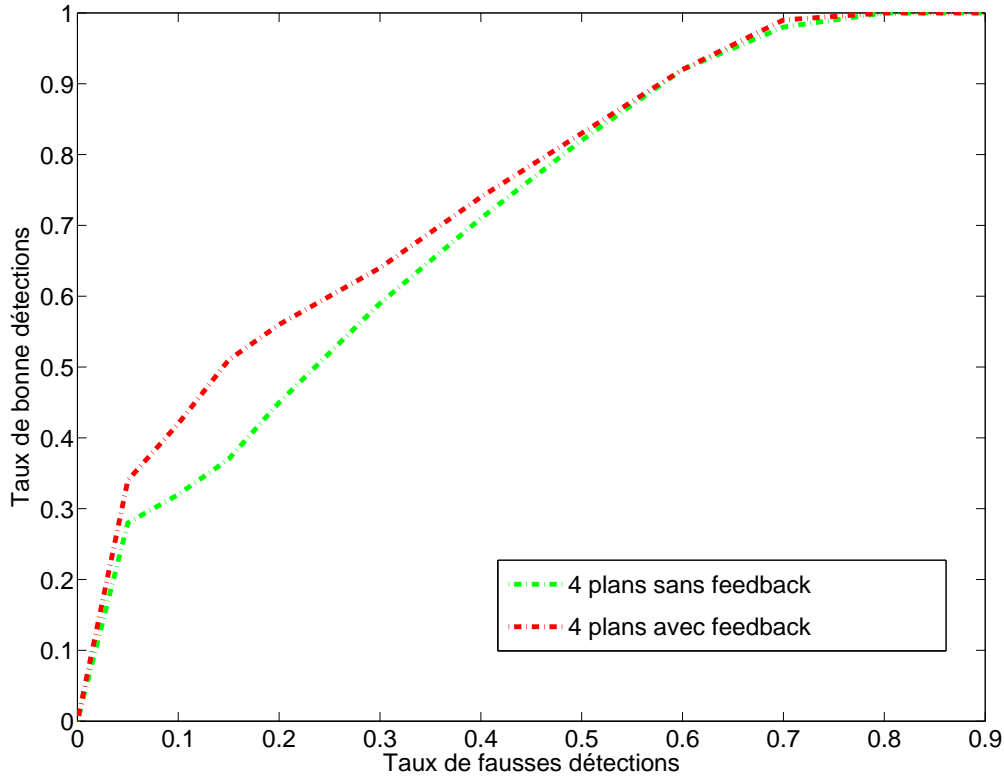


FIG. 5.35 – Superposition des 2 courbes ROC permettant de comparer les deux algorithmes de suivi (avec et sans *feedback*) sur le scénario **VIPA campus d'Aubière**.

en piste verte). En revanche, la second piéton situé au deuxième plan n'a pas été mieux suivi (passage de la piste bleu en piste rouge pour l'algorithme sans *feedback* et passage de la piste jaune en piste bleu ciel pour l'algorithme avec *feedback*). En effet, le piéton situé au second plan reste trop longtemps invisible dans l'image de profondeur pour permettre à la boucle de retour de l'associer avec une piste existante. Rappelons que le *feedback* permet de chercher dans l'image laser des amas de points qui n'ont pas été classifiés comme piéton par le module de détection quatre nappes. Il vise donc à aider l'algorithme de détection mais ne peut en aucun cas permettre de deviner ou émettre l'hypothèse qu'un piéton invisible dans l'image laser est bien toujours présent dans l'environnement observé.

Finalement et afin de démontrer l'avantage de l'algorithme de suivi avec *feedback* sur le module de suivi sans *feedback*, nous proposons un résultat de suivi temporel (cf. la figure 5.38) de ces deux algorithmes sur notre scénario vérité terrain **VIPA campus d'Aubière**. On remarque une amélioration nette de la qualité du suivi de l'algorithme avec *feedback* (en bleu) sur l'algorithme sans *feedback* (en rouge). En effet, l'algorithme avec *feedback* permet de mieux suivre la vérité terrain (en vert) que le module sans *feedback*.

## 5.4 Association de données centralisée laser/vidéo

L'objectif de cette section est de présenter les résultats obtenus par la méthode de fusion d'objets (fusion centralisée) provenant d'algorithmes de détection fondés sur la

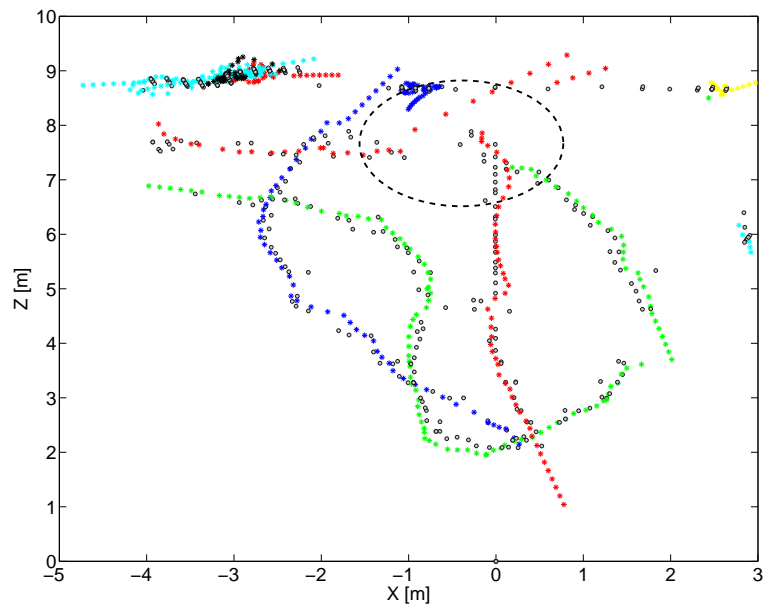


FIG. 5.36 – Résultats du suivi de deux piétons sur le scénario commenté dans la section 5.2.2 avec l’algorithme sans boucle de retour. Les mesures obtenues par l’algorithme de détection quatre plans sont représentées en gris. Les pistes sont identifiées par des étoiles de couleur.

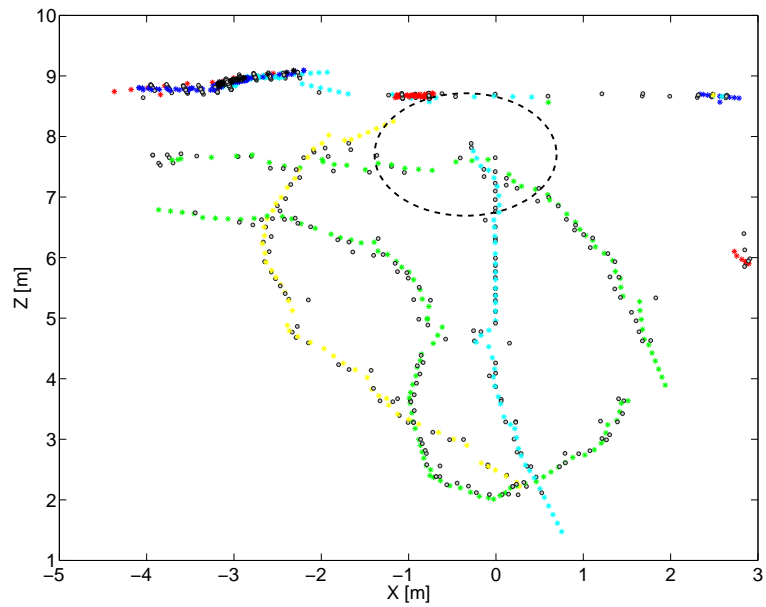


FIG. 5.37 – Résultats du suivi de deux piétons sur le scénario commenté dans la section 5.2.2 avec l’algorithme utilisant une boucle de retour. Les mesures obtenues par l’algorithme de détection quatre plans sont représentées en gris. Les pistes sont identifiées par des étoiles de couleur.

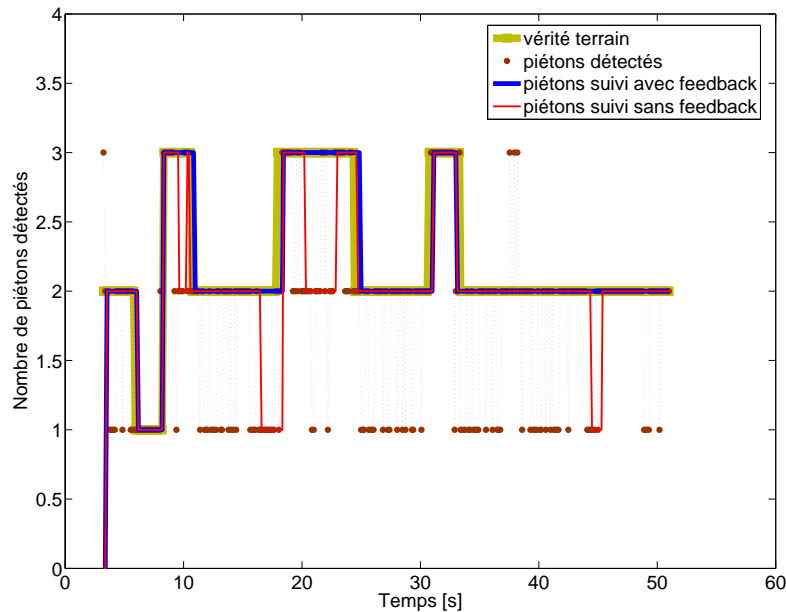


FIG. 5.38 – Résultats temporels des deux algorithmes de suivi (avec et sans *feedback*) sur le scénario **VIPA campus d'Aubière**.

combinaison de mesures fournies par un capteur laser et vidéo.

### 5.4.1 Stratégie et conséquences

A partir d'une image d'impacts laser, nous avons vu jusqu'à présent la difficulté de classifier avec certitude un obstacle comme étant un piéton car l'information délivrée est pauvre (amas de points), donc peu adaptée à la reconnaissance de piétons. En vision, la grande richesse d'information délivrée (information de couleur, de hauteur, de forme...) permet d'obtenir une meilleure reconnaissance d'« objets-piétons ». Pour cette raison, notre module de fusion centralisée donne une plus grande importance aux détections provenant d'un algorithme vidéo que pour celles fournies par un algorithme laser. Ce choix stratégique et ses conséquences sur le suivi des piétons méritent d'être explicitées maintenant :

- ▷ une piste suivie à partir de mesures uniquement vidéo survivra mais son facteur de confiance sur la fusion des données (*FCFD*) restera inférieur à 0.6.
- ▷ une piste suivie à partir de mesures uniquement laser sera détruite car le module considérera l'obstacle comme non piéton si la vidéo ne le reconnaît pas.
- ▷ une piste suivie à partir de mesures laser et vidéo survivra avec un facteur de confiance sur la fusion des données (*FCFD*) qui tendra vers 1.
- ▷ dans tous les autres cas, une piste suivie à partir de mesures laser et vidéo perturbées survivra mais son facteur de confiance sur la fusion des données (*FCFD*) variera entre 0.5 et 1. Cette information permet à l'utilisateur d'obtenir une information pertinente sur la qualité de fusion de la piste suivie.



### 5.4.2 Implémentation et séquences d'expérimentations

Le code est écrit en langage C++ et utilise la bibliothèque NT<sup>2</sup>. Etant donné que les mesures sont non synchronisées sur la fréquence d'acquisition d'un des deux capteurs (cf. le chapitre 4), deux vérités terrain ont été créées sur deux scénarios Renault. Les taux de bonnes et fausses détections sont alors fondés sur la fréquence d'acquisition des deux capteurs.

Le première séquence choisie propose de suivre un piéton qui traverse une route qui comporte des véhicules garés sur les côtés. La deuxième séquence choisie propose de se déplacer dans le centre-ville comportant une intersection et deux passages piétons. Plusieurs piétons (> 5) traversent la chaussée sur ces passages. De nombreux obstacles urbains tels que les véhicules, feux, poteaux, barrières, etc., sont sources de fausses alarmes. Les piétons détectés à partir de données télémétriques sont représentés par des points rouges tandis que ceux détectés à partir de données vidéos sont représentés par des rectangles bleus. Le résultat final du module de fusion centralisée est représenté par des rectangles jaunes.





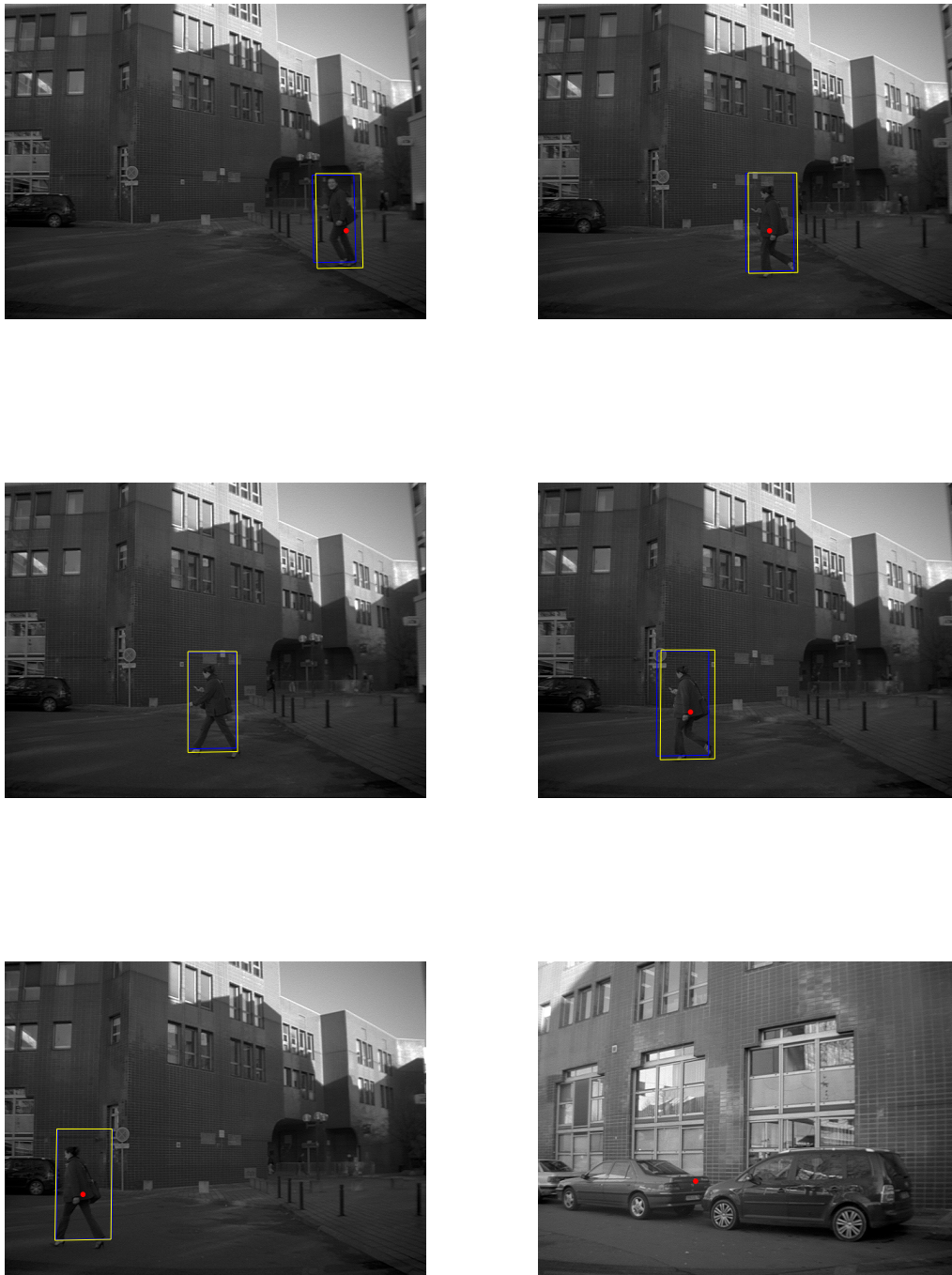


FIG. 5.39 – Séquence 1 : Résultats de l’algorithme de fusion centralisée sur un scénario livré par Renault.

Le tableau 1.3 montre l’avantage de la fusion centralisée pour diminuer significativement le nombre de fausses détections par rapport à l’utilisation seule des algorithmes basés laser ou vidéo. Notons également une amélioration du taux de détection de piéton après la fusion centralisée des données.

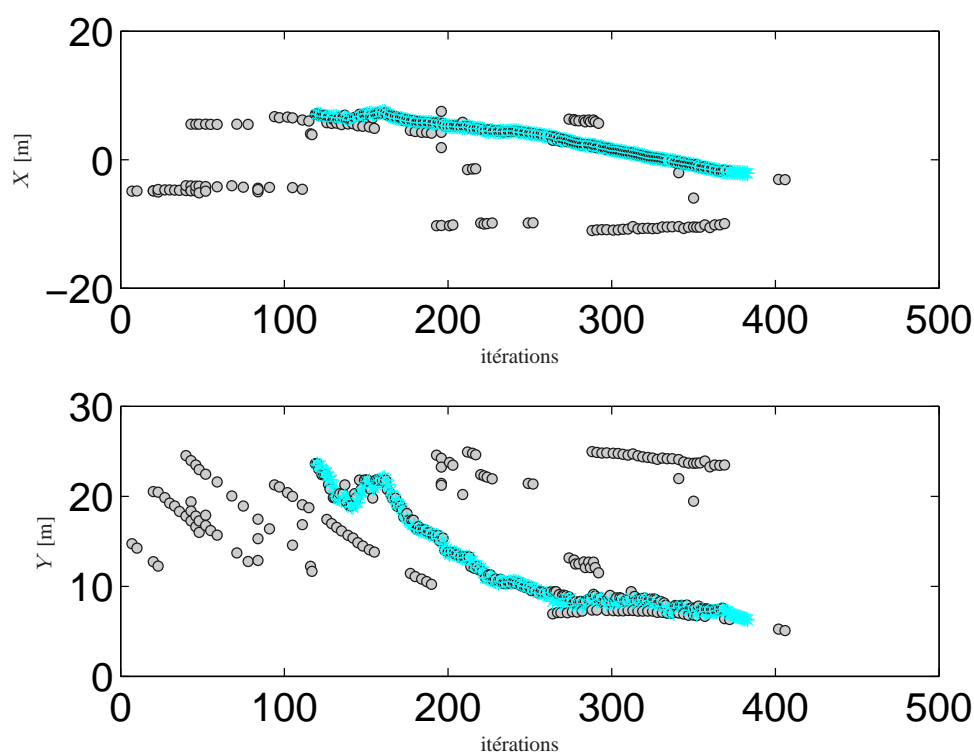


FIG. 5.40 – Séquence 1 : résultats du pistage en x et y : cas mono-cible d’une fusion centralisée laser-vidéo. Les mesures (vidéo et laser) sont représentées par des cercles gris. Chaque piste est représentée par une couleur différente.

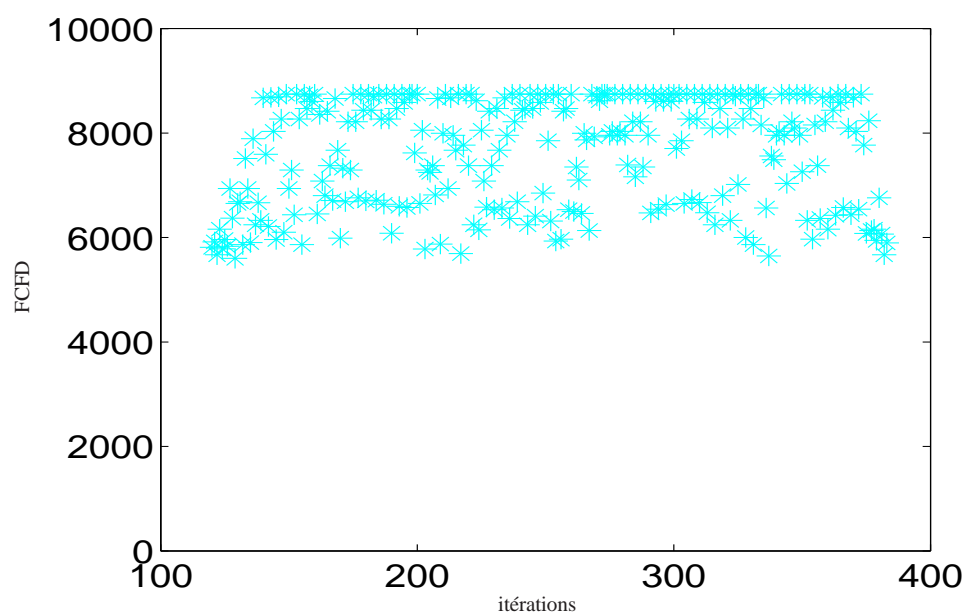
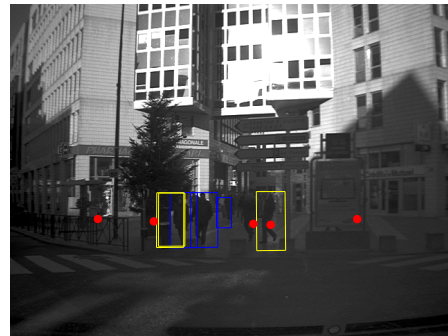
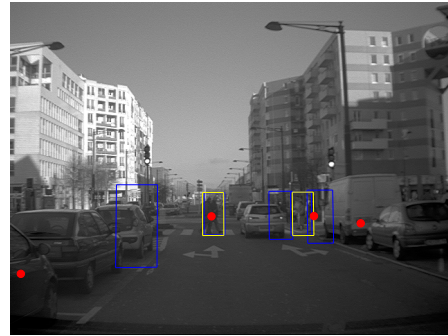
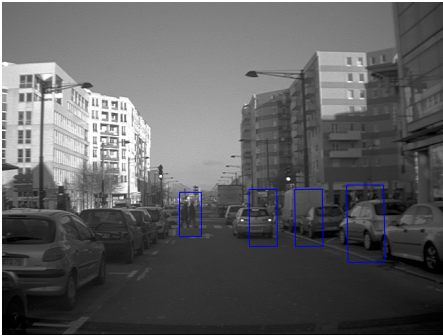


FIG. 5.41 – Séquence 1 : résultats du Facteur de Confiance sur la Fusion de Données (FCFD) associé à la figure 5.43. Chaque FCFD est représenté par une étoile de la même couleur que sa piste (cf. la figure 5.43).

TAB. 5.3 – Séquence 1 : taux de fausses et de bonnes détections sur scénario livré par Renault.

	taux de bonnes détections		taux de fausses détections	
	$\theta$	$\sigma_\theta$	$\theta$	$\sigma_\theta$
seulement laser	0.54	0.4974	0.3406	0.4463
seulement vidéo	0.8824	0.322	0.1929	0.2449
après fusion	0.9524	0.213	0.087	0.093



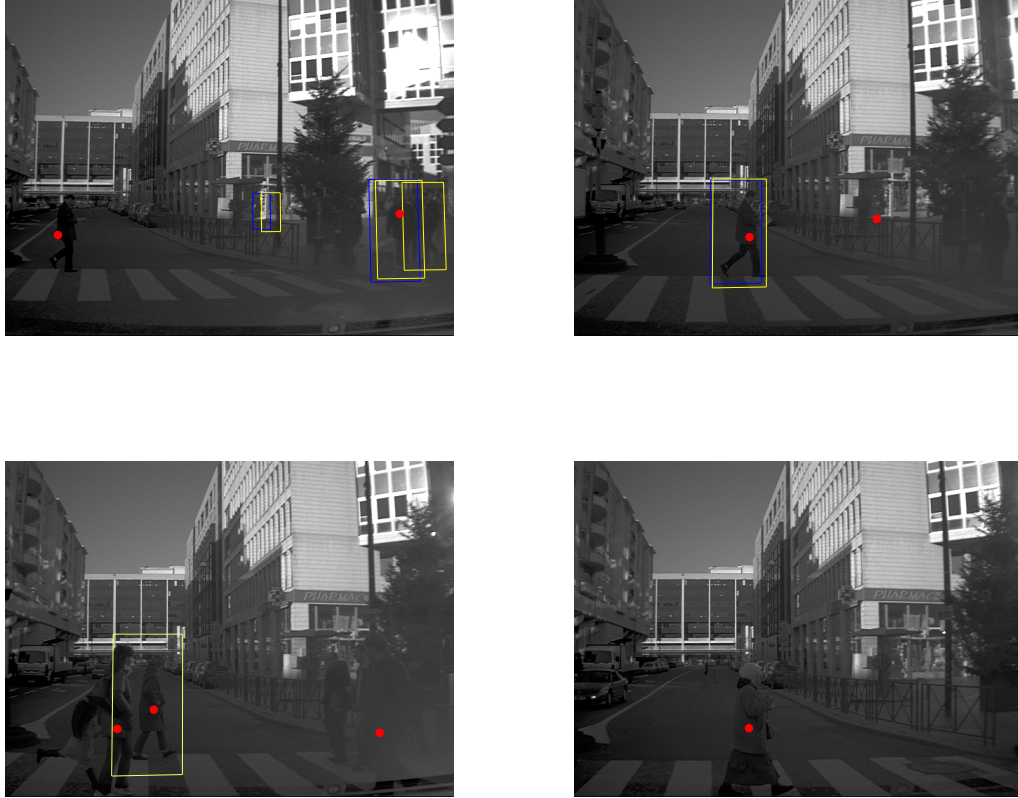


FIG. 5.42 – Séquence 2 : résultats de l’algorithme de fusion centralisée sur un scénario livré par Renault.

Le tableau 5.4 montre toujours le réel avantage de la fusion centralisée pour diminuer significativement le nombre de fausses détections par rapport à l’utilisation seule des algorithmes basés laser ou vidéo. En revanche le taux de détection de piétons après la fusion centralisée des données reste stable.

TAB. 5.4 – Séquence 2 : taux de fausses et de bonnes détections sur un scénario livré par Renault.

	taux de bonnes détections		taux de fausses détections	
	$\theta$	$\sigma_\theta$	$\theta$	$\sigma_\theta$
seulement laser	0.7289	0.3779	0.8367	0.3856
seulement vidéo	0.7191	0.4128	0.5962	0.4266
après fusion	0.708	0.3817	0.1864	0.2965

### 5.4.3 Synthèse : avantages, limitations et contraintes

Cette section a pour objectif d’illustrer l’apport mais aussi les limites liées à l’exploitation d’une fusion centralisée laser/vidéo. Quelques images-clés associées à ces résultats qualitatifs (tableaux, courbes) sur deux séquences Renault ont pu en donner une idée.

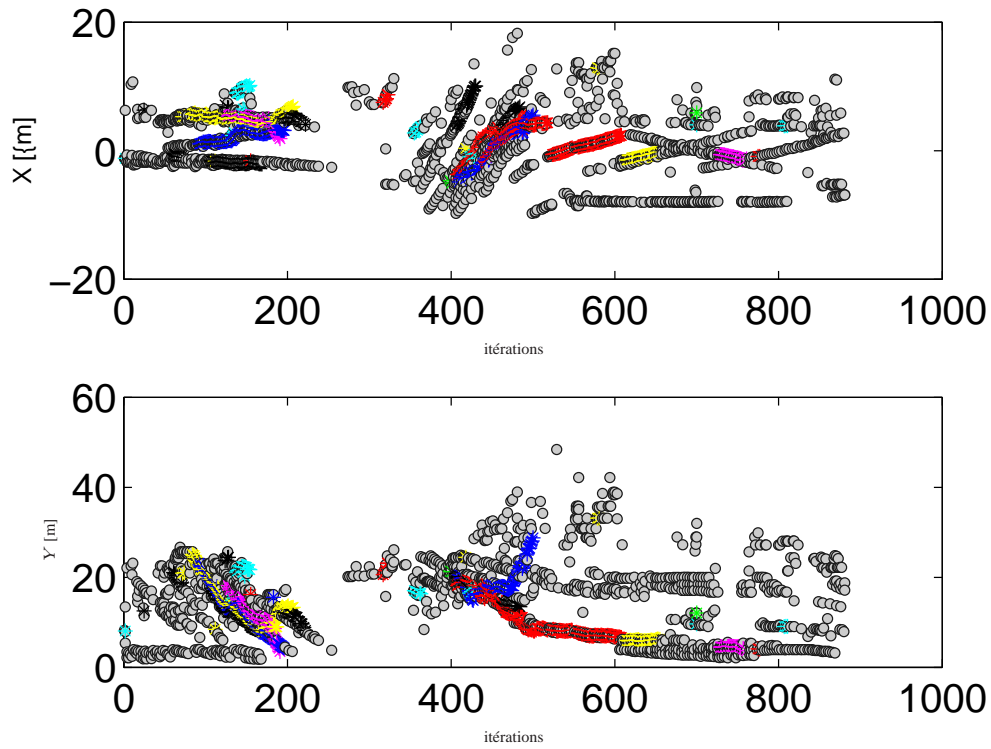


FIG. 5.43 – Séquence 2 : résultats du pistage en x et y : cas mono-cible d’une fusion centralisée laser-véo. Les mesures (vidéo et laser) sont représentées par des cercles gris. Chaque piste est représentée par une couleur différente.

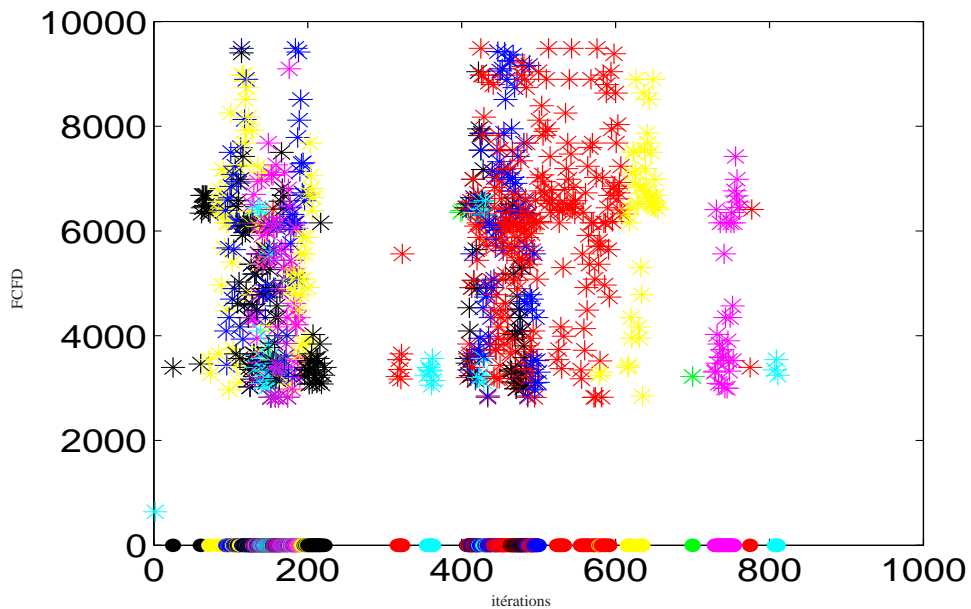


FIG. 5.44 – Séquence 2 : résultats du Facteur de Confiance sur la Fusion de Données (FCFD) associé à la figure 5.43. Chaque FCFD est représenté par une étoile de la même couleur que sa piste (cf. la figure 5.43).



Il est clairement illustré par la figure 5.39 que la combinaison centralisée des mesures délivrées par deux capteurs (laser et vidéo) permet de filtrer efficacement les fausses alarmes émises par l'algorithme de détection laser. Les résultats du tableau 5.3 démontrent une nette amélioration des taux de bonnes et de fausses détections après la fusion centralisée des données. En revanche, si la séquence illustrée en figure 5.39 fonctionne bien, cela est moins concluant sur la séquence illustrée en figure 5.42. Le tableau 5.4 montre toujours une nette amélioration du taux de fausses détections après la fusion centralisée mais le nombre de piétons suivis reste stable.

En effet, pour suivre les piétons correctement avec le module de fusion centralisée proposé dans ce document, il faut remplir au moins deux conditions : 1) disposer d'un détecteur de piétons vidéo efficace, et 2) disposer de détection laser et vidéo homogène en position, c'est-à-dire permettant de pouvoir associer les différentes mesures entre elles.

Si la première condition est facile pour des applications de détection de piéton non contraintes par le temps réel, elle devient beaucoup plus délicate lorsque cette contrainte est nécessaire. La deuxième condition est difficile à obtenir à partir de mesures fournies par un télémètre et une caméra. En effet, si l'estimation de profondeur est facile à obtenir avec précision avec un laser, il n'en est pas de même avec une caméra 2D qui peut délivrer des erreurs de plusieurs mètres rendant difficile l'association entre les mesures laser et vidéo.

Comme dans la séquence 2, nous ne remplissons aucune de ces deux conditions, les performances de la fusion pour le suivi des piétons deviennent équivalentes à l'utilisation séparée des algorithmes de détection mono-capteur.

Enfin, l'installation d'un système multi-capteurs reposant sur un télémètre et une caméra impose une contrainte de calibration extrinsèque permettant par exemple de pouvoir projeter les mesures laser dans l'image vidéo.

## 5.5 Conclusion

Nous avons dédié ce chapitre à la validation des différents algorithmes sur de nombreux scénarios de données réelles qui peuvent être classés en deux catégories principales :

- ▷ ceux acquis à bord du véhicule expérimental du LASMEA proposant divers types d'environnements tels que l'intérieur d'un bâtiment ou le campus d'Aubière, à des vitesses comprises pour des raisons de sécurité entre 0 et 7 *km/h*,
- ▷ et ceux acquis à bord du véhicule expérimental de Renault proposant des environnements urbains complexes tels que les intersections, et différentes voies de circulations à des vitesses comprises entre 0 et 50 *km/h*.

Dans un premier temps, nous avons présenté les deux scénarios vérité terrain qui nous ont permis de comparer les performances des algorithmes de détections. La construction de tels scénarios est très difficile dans le cas général (de manière automatique) où la position des piétons environnant est inconnue. La technique retenue (obtention manuelle), constitue dans notre cas la meilleure méthode car elle permet de générer une vérité terrain « à la carte » c'est-à-dire permettant à l'utilisateur de décider à chaque image quel piéton doit être détecté ou pas en fonction des spécifications qu'il doit suivre.

Dans un second temps, les résultats très satisfaisants obtenus par l'algorithme de détection fondé sur la combinaison des quatre plans laser ont été présentés. Une comparaison systématique entre l'algorithme quatre nappes et l'algorithme mono-nappe sur toutes les

images vidéos de la section a permis de démontrer la supériorité de l'algorithme quatre nappes sur l'algorithme mono-nappe lorsque les détections situées dans les quatre nappes laser sont combinées judicieusement avant de prendre une décision finale. En effet, à cause du tangage du véhicule ou du problème de l'occultation des objets entre eux, aucune nappe ne délivre les mêmes informations de la scène située à l'avant du véhicule. Ce travail a démontré que la combinaison des quatre plans laser donne une réponse à ces problèmes en permettant de mieux détecter les piétons et en diminuant de manière significative le nombre de fausses détections.

Ensuite deux méthodes de suivi ont été proposées (l'une étant l'amélioration de l'autre). La première méthode repose sur le schéma classique qui décorrèle complètement la partie détection de la partie suivi. Les expérimentations présentées dans la section 5.2.2 ont montré que les piétons non localisés par l'algorithme de détection perturbe l'algorithme de suivi qui n'assure plus la cohérence temporelle des pistes. La seconde méthode propose donc de répondre à ce problème en ajoutant un *feedback* entre la sortie du module de suivi et l'image laser. Cela revient à corrélér le module de suivi avec le module de détection en les rendant complémentaires. Les nombreux tests sur des séquences réelles permettent d'attester de la pertinence de la méthode. En effet, des trajectoires subissant des occultations plus ou moins importantes dans des environnements urbains complexes ont été reconstruites avec succès.

Enfin, nous avons fini ce chapitre par la présentation des résultats obtenus par l'algorithme de fusion centralisée provenant d'algorithmes de détection mono-capteur laser et vidéo. Les résultats obtenus sont mitigés. En effet, une stratégie donnant une grande légitimité aux mesures vidéos rend le module de fusion dépendant des performances du module de vision. Il existe donc une marge de progrès sur le procédé de fusion des observations comme combiner par exemple les architectures de fusions série et centralisée [Premibida *et al.*, 2007] [Ngako-Pangop *et al.*, 2008] permettant de mieux prendre en compte les mesures laser. Nous travaillons donc actuellement sur un algorithme de fusion dit « série » entre un capteur laser et un capteur vidéo. L'idée est d'utiliser les détections laser comme fenêtre d'intérêt dans l'image vidéo. Ces détections sont ensuite analysées par un algorithme de reconnaissance de forme comme AdaBoost par exemple. La vision est donc utilisée pour confirmer ou infirmer les objets proposés par le laser. Les deux capteurs ne sont alors plus traités de manière parallèle comme dans l'architecture développée dans ce document mais selon un enchaînement dit « en série » .



---

## Conclusion générale et perspectives

Dans ce mémoire, nous nous sommes intéressés au problème de détection et de suivi de piéton à bord d'un véhicule routier circulant dans un environnement urbain complexe.

Si pour un être humain il est aisé de reconnaître un piéton au milieu de la circulation urbaine, une telle reconnaissance n'est pas facile pour un système de perception artificielle basé sur des capteurs comme les caméras, lasers, radars, et autres. L'utilisation des capteurs vidéo pour résoudre le problème de détection et d'identification des piétons semble au premier abord naturelle de par leur capacité à pouvoir saisir la forme et la texture d'un objet. Cependant, un algorithme de reconnaissance doit examiner un très grand nombre de régions ce qui nécessite un coût de calcul excessif pour un algorithme temps-réel. De plus, l'impossibilité d'obtenir directement l'information de profondeur avec précision nous a, entre autres raisons, fait pencher vers la mise au point d'une méthode de détection à partir d'un capteur laser.

Afin de caractériser au mieux la trajectoire complexe d'un piéton, nous nous sommes intéressés à une modélisation probabiliste fondée sur les filtres bayésiens. En effet, ces algorithmes ont su démontrer leur robustesse dans de nombreuses applications telles que la navigation, le suivi d'entités en vision par ordinateur, la robotique mobile, etc. Ces algorithmes nécessitent la définition d'un modèle d'évolution du système dynamique et d'un modèle de vraisemblance des données.

Le choix de l'algorithme de filtrage doit être fonction de la nature de l'entité suivie. En effet, dans le cas d'un modèle dynamique linéaire à bruit gaussien, le filtre de Kalman fournit la meilleure solution au problème. En revanche, devant les trajectoires complexes et imprévisibles d'un piéton, l'utilisation des méthodes séquentielles de Monte Carlo (filtre particulaire) apporte une solution satisfaisante au problème.

Enfin, il est évident qu'une détection de piéton fondée sur un unique capteur laser ne peut donner à l'heure actuelle pleinement satisfaction car trouver des situations de mise en défaut est chose relativement aisée. En effet, la pauvreté de l'information délivrée par ce type de capteur (aucune information sur la hauteur, les contours, la couleur des objets) sur un environnement aussi complexe et varié que les voies de circulation d'une ville nous a amené à investiguer une solution multisensorielle laser-vidéo. Cette stratégie est apparue comme une bonne solution de par sa complémentarité : le capteur laser détecte naturellement la position en profondeur des objets tandis que le capteur vidéo apporte des informations sur la forme, la texture et la couleur de ces objets.

## Synthèse des travaux effectués

Ces travaux ont consisté à proposer un algorithme de détection et de suivi de piéton à partir d'un unique capteur laser quatre plans puis d'une collaboration entre un capteur

laser et une caméra vidéo. La validation expérimentale de ces algorithmes a été menée sur de nombreux scénarios plus ou moins complexes à bord, d'une part du véhicule expérimental du LASMEA et d'autre part, d'un véhicule expérimental appartenant à Renault.

### **Proposition d'un algorithme de détection de piéton à partir d'un capteur laser quatre plans**

Afin de disposer d'un algorithme de détection robuste de piéton en environnement urbain à partir d'un capteur laser, nous nous sommes intéressés au développement d'une méthode reposant sur un capteur laser quatre plans tel que ceux développés par la société IBEO. Notre approche repose sur l'idée de fusionner (combiner) toute l'information localisée dans les quatre nappes laser avant de prendre une décision finale. En effet, et afin d'améliorer les performances d'un algorithme de détection mono-nappe, l'idée est de compenser une perte momentanée d'informations localisées dans une nappe, due par exemple au tangage du véhicule, par les autres nappes laser.

#### **Contributions**

- ▷ Calcul d'une fonction de vraisemblance relative à la distance entre les points d'une image de profondeur permettant d'extraire des « objets piétons »,
- ▷ Calcul d'une combinaison décentralisée des informations situés dans les différents plans laser pour améliorer les performances d'un algorithme mono-plan.

#### **Résultats**

Une première campagne d'expérimentations nous a permis d'évaluer l'efficacité de notre méthode d'extraction d'« objets piétons » à partir d'une image de profondeur. Par la suite, les nombreux résultats menés à bord des véhicules du LASMEA et de Renault ont montré que l'utilisation judicieuse des quatre plans laser permet de mieux détecter les piétons et de diminuer sensiblement le nombre de fausses détections.

### **Proposition d'un algorithme de suivi avec l'ajout d'un feedback entre le module de suivi et l'image laser**

Nous avons présenté dans ce mémoire un filtre de suivi multi-objets générique reposant sur l'utilisation d'un filtre à particules afin de caractériser au mieux la trajectoire complexe d'un piéton. Classiquement, le module de suivi est placé après le module de détection ce qui le rend complètement dépendant des performances du module de détection. En effet, dans le cas où le module de détection manque certains piétons pour diverses raisons (occultation partielle ou totale, vêtement sombre, mauvaise condition atmosphérique, etc.), le module de suivi n'a pas de mesures à associer à la piste piéton et risque de la détruire.

Afin de remédier à ce problème, nous avons décidé d'insérer un feedback entre le module de suivi et le module de détection permettant au premier nommé d'aider et de guider le module de détection dans sa recherche de piétons lors des cas litigieux et plus compliqués. Les résultats obtenus sur des données réelles ont montré l'amélioration du taux de détection apportée par la méthode.

### Contributions

- ▷ Filtre particulière SIR basé sur un module d'association de données KDE permettant de mieux tenir compte du caractère non gaussien du nuage de particules lors de l'association mesure-objet,
- ▷ Filtre avec une boucle de retour dans l'image laser permettant au module de suivi de se renseigner directement dans l'image laser lorsque le module de détection ne lui fournit pas le résultat attendu.

### Résultats

Les résultats obtenus sur des données réelles montrent l'intérêt du *feedback* entre les modules de suivi et de détection pour améliorer le taux de détection globale. Le filtre, très générique, est capable de suivre des piétons dans de nombreuses situations telles que des personnes qui marchent, courent ou changent brusquement de direction.

### Proposition d'un algorithme de fusion laser-vidéo

Comme l'attestent les résultats actuels, obtenir un détecteur robuste de piétons à partir d'un algorithme mono-capteur reste quelque chose de difficile surtout en situation routière urbaine. En effet, chaque capteur possède des avantages mais aussi des inconvénients. La fusion de données vise à associer et/ou combiner les avantages de ces différents capteurs et de minimiser au maximum leurs inconvénients, dans le but d'obtenir une meilleure information par rapport à l'utilisation séparée de ces différentes sources de données. Une architecture de fusion centralisée permettant de combiner des données provenant d'un capteur laser et d'une caméra vidéo a été proposée. Cet algorithme repose sur le même formalisme fondé sur les filtres bayésiens récurrents avec une association de données non paramétrique afin de mieux prendre en compte toute l'information disponible, et notamment, la note de détection globale (NDG) d'une détection.

### Contributions

- ▷ Modélisation de la fonction de vraisemblance par le mélange d'une loi gaussienne avec une loi uniforme afin de mieux combiner l'information fournie par les différents modules de détection mono-capteur,
- ▷ Calcul d'une note de fusion, permettant de renseigner sur la confiance que l'algorithme de fusion donne à la piste laser-vidéo.

### Résultats

Les expérimentations présentées dans le chapitre 5 montrent la pertinence de cette méthode afin de diminuer significativement le nombre de fausses détections fournies par les algorithmes de détection mono-capteur. En effet, de nombreuses fausses détections sont naturellement filtrées par l'étape de fusion lorsque les deux capteurs ne détectent pas les mêmes objets aux mêmes endroits et au même moment.

### Perspectives

Le système présenté dans ce mémoire est une méthode générique qui peut être appliquée à d'autres technologies de perception et dont les performances peuvent être grandement améliorées par les travaux suivants :

#### Amélioration du module de détection laser

Nous avons vu au chapitre 4 que pour chaque scan, le module de détection commence par filtrer tous les amas de points présentant une taille géométrique importante et qui peuvent être relatifs à des murs, voitures ou panneaux afin de ne garder que les « petits » objets qui peuvent être assimilés à un piéton. Ce choix pose problème lorsque les objets sont mal filtrés, car ils peuvent générer des détections « fantômes ».

Pour éviter ce phénomène, une alternative est de ne plus filtrer préalablement l'image de profondeur, afin de pouvoir revenir par la suite vérifier si une détection classée piéton par la brique de reconnaissance est vraie. Cette stratégie doit permettre d'éviter de classer comme piéton, des amas de points appartenant à une voiture ou à un mur mal filtré par exemple.

#### Développement d'un module de détection basé radar

L'algorithme développé pour détecter des piétons à partir d'une image laser pourrait être adapté à la détection de piéton à partir d'une image radar car ce capteur possède deux atouts majeurs :

- ▷ il peut acquérir des images de bonne qualité par des conditions atmosphériques difficiles comme la brume, le brouillard, ou la pluie,
- ▷ il peut fournir directement l'information de position et l'information de vitesse des cibles détectées..

De plus, ces capteurs sont déjà implantés sur les véhicules automobiles « hautes gammes ». Développer un détecteur de piéton à partir d'un tel capteur, intéresserait donc grandement les constructeurs car il permettrait de diminuer les coût de développement.

#### Evaluation des résultats de détection

De manière générale, l'évaluation d'une méthode de détection fondée sur un capteur laser reste problématique. Typiquement, il n'existe aucune base de données laser qui soit reconnue par la communauté scientifique pour tester avec les mêmes critères, les performances des algorithmes de détection fondés sur un télémètre laser. Pour le moment, chacun valide ses résultats sur sa propre base de données rendant quasiment impossible une comparaison intéressante des principales méthodes existantes. Proposer une telle base permettrait une réelle avancée sur l'évaluation des performances de chaque algorithme.

#### Approfondir le module de suivi

Les méthodes de suivi multi-cibles ont fait l'objet de nombreux travaux ces dernières années. Pourtant, en raison de problèmes encore ouverts, elles constituent toujours une voie d'étude à explorer. Mentionnons quelques difficultés parmi d'autres comme le nombre variable d'objets à suivre, l'occultation ou le croisement d'un objet par un autre

sont encore des points durs qui mettent un grand nombre de méthodes en défaut. Proposer de nouvelles méthodes de combinaison de l'information situées dans chaque particule avec l'introduction de la notion d'occultation pourrait améliorer la capacité de suivi de l'algorithme de pistage.

### **Développement d'un nouveau module de fusion**

Dans cette thèse, les deux algorithmes de fusion reposent sur une méthode parallèle de combinaison de l'information. Même si les résultats obtenus dans ce document sont satisfaisants, il semble dommage de ne pas se servir des détections laser afin de guider la reconnaissance vidéo qui s'en retrouvera du coup simplifiée mais surtout plus efficace. Le développement d'un module de « fusion série » combiné à un module de fusion centralisée, se pose comme une suite immédiate des travaux évoqués dans ce document.



---

## Annexe A

# Bilan représentatif des principaux capteurs laser proposés sur le marché




Cette annexe présente une sélection représentative des capteurs laser disponibles sur le marché. Elle est renseignée sous forme d'une base de données permettant de comparer les critères suivant :

- ▷ portée : distance maximale de détection du capteur,
- ▷ champ : angle d'ouverture maximal de détection du télémètre,
- ▷ résolution : nombre de points maximal constituant l'image laser.

Dans ce tableau, on peut recenser trois grandes familles de constructeur :

- ▷ les capteurs mono-nappe de chez SICK,
- ▷ les capteurs multi-nappes de chez IBEO,
- ▷ les imageurs 3D de chez Riegl.

Enfin, le HDL-64E de Velodyne <sup>1</sup> est un produit récent qui semble être apparemment « au dessus du lot ». Les premiers retours d'expérience du projet ANR CityVIP confirment la qualité et le potentiel du produit.

Fabricant - Référence	Portée	Champ	Résolution	Aspect
Sick LMS 221	80 m	180°	0.25°	
Sick LDA	100 m	360°	0.125°	
Sick NAV 200	30 m	360°	0.1°	

---

<sup>1</sup><http://www.velodyne.com/lidar/>



A. BILAN REPRÉSENTATIF DES PRINCIPAUX CAPTEURS LASER PROPOSÉS SUR LE  
MARCHÉ

Fabricant - Référence	Portée	Champ	Résolution	Aspect
HOKUYO UTM-30LX	30 m	270°	0.25°	
IBEO LD ML	80 m	270°*3.2°	0.25°*1.07°	
IBEO Alasca	128 m	240°*3.2°	0.25°*1.07°	
Riegl LMS-Q120	150 m	80°	0.01°	
Riegl LMSZ210-i	400 m	360°*80°	0.005°*0.005°	
Riegl LMSZ210-60	120 m	333°*60°	0.018°*0.036°	
Velodyne HDL-64E	120 m	360°*26.8°	0.09°*0.4°	
ZF LASER IMAGEUR 6006	79 m	360°*310°	0.0016°*0.0016°	

---

## Annexe B

# Modèles d'évolution pour la trajectographie d'entités piéton

La mise en oeuvre d'un algorithme de pistage suppose une bonne description du modèle d'évolution  $p(\mathbf{X}_k|\mathbf{X}_{k-1})$  du système à estimer. La nature du modèle dépend directement du type d'entité suivie et de la connaissance que nous avons de cette entité. En effet, quel que soit l'objet suivi, l'objectif est de construire les trajectoires spatio-temporelles de ces objets. Ce modèle doit permettre de prédire leur évolution dans le temps à tout instant et donc d'anticiper sur des actions à entreprendre ou le cas échéant, de compenser des périodes de croisement et d'occultation. Dans notre cas, les objets que nous souhaitons suivre dans le temps sont des piétons. Devant la complexité du déplacement d'un piéton pouvant comporter des changements soudain de direction, aucun modèle d'évolution ne s'est imposé dans la littérature même récente. Cette annexe propose d'apporter quelques pistes.

### B.1 Etude du déplacement d'un piéton

Dans ces travaux portant sur la modélisation de mouvement des piétons, Tecknomo [Teknomo, 2002] établit une distinction entre le déplacement d'un piéton « isolé » avec ceux membres d'un groupe. Dans le déplacement en groupe, les mouvements des différents piétons sont agrégés au mouvement du groupe dans sa globalité ce qui simplifie la modélisation de leurs trajectoires par une modélisation linéaire car un groupe ne peut pas changer subitement de direction. A l'inverse d'un piéton seul, son comportement ne peut pas être connu à l'avance de façon certaine car il peut changer à tout moment de direction ce qui rend très difficile la modélisation de ses trajectoires. Dans le cadre de nos travaux, ce sont les piétons « isolés » qui nous intéressent et retiennent notre attention car la majorité des accidents véhicules piétons (plus de 95%) n'impliquent qu'un seul piéton à la fois [obs, 1991].

A partir de ce qui vient d'être dit sur l'étude comportementale des piétons, une loi d'évolution doit être proposée.

#### B.1.1 Modèle d'évolution à vitesse constante

Lorsqu'une modélisation physique du mouvement de l'objet est disponible, la loi d'évolution est connue sous forme d'un système d'équation linéaire. Un exemple typique

est l'utilisation d'un modèle à vitesse constante comme celui utilisé par Blanc [Blanc, 2005] dans des applications de suivi de véhicule routier.

Le modèle d'évolution peut être présenté sous forme matricielle par :

$$\mathbf{X}_k = \mathbf{F}\mathbf{X}_{k-1} + \mathbf{G}\mathbf{v}_k \quad (\text{B.1})$$

Le modèle à vitesse constante est donné par les matrices suivantes :

$$\mathbf{F}_k = \begin{bmatrix} 1 & \Delta T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \frac{\Delta T^2}{2} & 0 \\ \Delta T & 0 \\ 0 & \frac{\Delta T^2}{2} \\ 0 & \Delta T \end{bmatrix} \quad (\text{B.2})$$

où  $\mathbf{X}_k = (\mathbf{x}, \dot{\mathbf{x}}, \mathbf{y}, \dot{\mathbf{y}})$  est le vecteur d'état,  $\mathbf{F}_k$  la matrice de transition qui modélise l'évolution de  $\mathbf{X}_k$  et  $\mathbf{G}$  la matrice de bruit sur le modèle considéré ici comme étant l'accélération. Le bruit d'évolution  $\mathbf{v}_k$  est supposé suivre une fonction gaussienne de moyenne nulle et de covariance  $\mathbf{Q}_k$  tel que  $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$ .

$$\mathbf{Q}_k = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \quad (\text{B.3})$$

### B.1.2 Modèle d'évolution à mouvement circulaire

On peut également mentionner les modèles d'évolutions cités dans [Li & Jilkov, 2000] pour décrire la dynamique des mouvements d'un piéton. Notamment, il est possible de mieux prédire les changements possibles de trajectoires par un modèle d'évolution à mouvement circulaire [Nabaa & Bishop, 1999] où l'angle de cap  $\theta_k$  est utilisé comme perturbation de la trajectoire prédite. Le modèle à mouvement circulaire appliqué est défini ci-dessous :

$$\mathbf{X}_k = \mathbf{F}\mathbf{X}_{k-1} + \mathbf{v}_k \quad (\text{B.4})$$

avec

$$\mathbf{F}_k = \begin{bmatrix} 1 & \frac{\Delta T \sin(\theta_k)}{\theta_k} & 0 & -\frac{\Delta T}{\theta_k} \cdot (1 - \cos(\theta_k)) \\ 0 & \cos(\theta_k) & 0 & -\sin(\theta_k) \\ 0 & \frac{\Delta T}{\theta_k} \cdot (1 - \cos(\theta_k)) & 1 & -\frac{\Delta T \sin(\theta_k)}{\theta_k} \\ 0 & \sin(\theta_k) & 0 & \cos(\theta_k) \end{bmatrix} \quad (\text{B.5})$$

où  $\mathbf{X}_k = (\mathbf{x}, \dot{\mathbf{x}}, \mathbf{y}, \dot{\mathbf{y}})$  est le vecteur d'état,  $\mathbf{F}_k$  la matrice de transition qui modélise l'évolution de  $\mathbf{X}_k$ , et  $\theta_{k+1} = \theta_k + b_g$  avec  $b_g \sim \mathcal{N}(0, \sigma_g)$ .  $\sigma_g$  est l'écart-type du bruit sur le cap de la trajectoire du piéton.

### Modèle de vraisemblance

L'équation de vraisemblance  $p(\mathbf{Z}_k | \mathbf{X}_k)$  décrit le lien entre une mesure  $\mathbf{Z}_k$  et une hypothèse de l'état courant des objets  $\mathbf{X}_k$ .

Dans ces simulations, l'état recherché est la position du piéton. La mesure fournie par le module de détection correspond donc à la meilleure position candidate estimée. L'équation de vraisemblance (B.6) permettant de modéliser l'imprécision des mesures est définie par :

$$P(\mathbf{Z}_k|\mathbf{X}_k) = \frac{1}{2\pi \cdot |\mathbf{R}|^{1/2}} \exp \left[ -\frac{1}{2}(\mathbf{Z}_k - \mathbf{X}_k)^t \mathbf{R}^{-1}(\mathbf{Z}_k - \mathbf{X}_k) \right] \quad (\text{B.6})$$

La matrice  $\mathbf{R}$  représente le bruit sur la mesure définie par :

$$\mathbf{R} = \begin{bmatrix} \sigma_{cx}^2 & 0 \\ 0 & \sigma_{cy}^2 \end{bmatrix} \quad (\text{B.7})$$

L'équation de vraisemblance pondère les nouvelles particules par la vraisemblance de l'observation  $\mathbf{Z}_k$  à l'instant  $k$ , sachant l'échantillon d'état  $\mathbf{X}_k$  :  $\pi_k = P(\mathbf{Z}_k|\mathbf{X}_k)$ .

## B.2 Evaluation des trajectoires et comparaisons

### B.2.1 Simulations de trajectoires

Dans cette section, nous proposons d'évaluer la pertinence du suivi des deux modèles d'évolutions présentés dans cette annexe :

- ▷ modèle à vitesse constante,
- ▷ modèle à mouvement circulaire.

Pour comparer ces deux modèles, plusieurs trajectoires comportant différentes difficultés telles que des lignes droites, des courbes et des changements brusques de direction sont simulées.

Elles sont analysées selon trois critères :

- ▷ la qualité du suivi en position du piéton.
- ▷ la qualité de l'estimation de la vitesse de déplacement du piéton.
- ▷ l'efficacité de l'échantillonnage. Ce critère supplémentaire permet de montrer la capacité d'un modèle d'évolution à maintenir une bonne cohérence temporelle comme lors d'un changement brusque de rythme, en convergeant rapidement vers le nouvel état de la cible.

#### Efficacité de l'échantillonnage

Pour quantifier le taux de particules efficaces d'une population de  $N$  particules, affectées de leurs poids  $\{w_k^n\}_{n=1}^N$  normalisés tels que  $\sum_{n=1}^N w_k^n = 1$ , nous utilisons le critère proposé dans [Gordon *et al.*, 1993b] :

$$N_{eff} = \frac{1}{N \sum_{n=1}^N (w_k^n)^2} \quad (\text{B.8})$$

Ce critère tend vers 0 lorsque très peu de particules ont de gros poids, c'est-à-dire lorsque la région intéressante de l'espace d'état est sous-échantillonnée. Il tend vers 1 lorsque toutes les particules ont des poids à peu près identiques, c'est-à-dire lorsqu'elles ont réussi à peupler la région intéressante de l'espace d'état.

#### Evaluation de l'erreur en suivi

La précision du suivi peut être quantifiée par la moyenne du cumul des distances Euclidiennes entre la position des particules estimées  $\hat{p}(k, n)$  de chaque objet, et la position « vérité terrain »  $p_{VT}(k)$ .

$$E(k) = \frac{1}{N} \sum_{n=1}^N \sqrt{(p_{VT}(k) - \hat{p}(k, n))^T (p_{VT}(k) - \hat{p}(k, n))} \quad (\text{B.9})$$

avec  $N$  le nombre de particules du filtre.

## B.2.2 Résultats

Le suivi a été mené sur deux scénarios :

- ▷ le premier propose de suivre un piéton sur une trajectoire linéaire simple ne comportant aucun changement brusque de direction et de rythme,
- ▷ le deuxième propose de suivre un piéton sur une trajectoire complexe comportant un changement brusque de direction.

### Modèle à vitesse constante

Pour ces simulations, le nombre de particules du filtre SIR est fixé à  $N = 400$ . La dynamique est définie par le module à vitesse constante selon les équations B.1, B.2 et B.3 avec  $\sigma_x = \sigma_y = 2m.s^{-2}$ . Le bruit de mesure

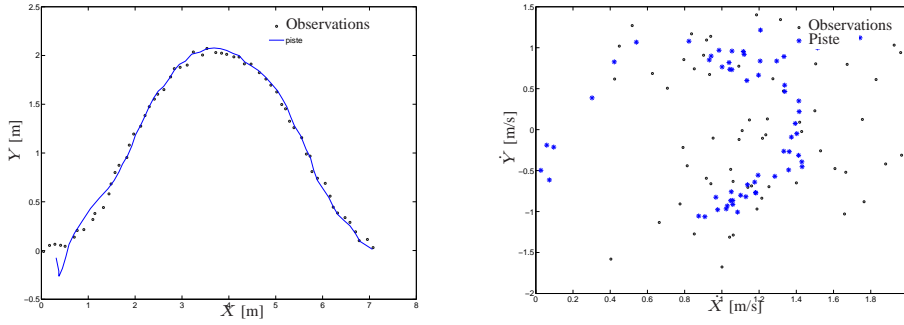


FIG. B.1 – Scénario qui propose le suivi d'un piéton sur la trajectoire en forme de courbe. A gauche, vue de dessus du résultat de suivi en position. A droite, vue de dessus du résultat de suivi en vitesse. Les mesures sont représentées par des cercles gris.

## B.2. EVALUATION DES TRAJECTOIRES ET COMPARAISONS

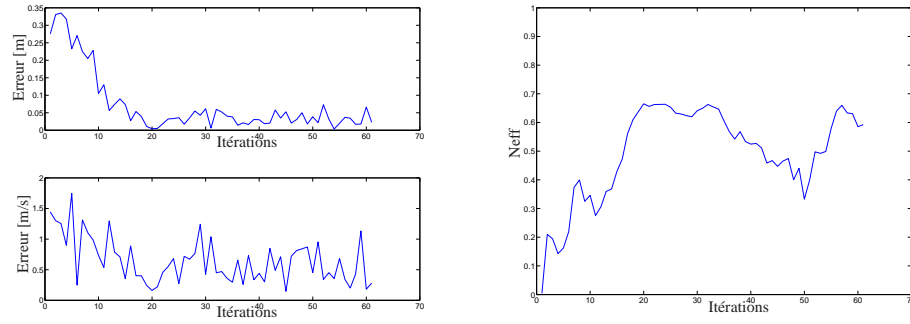


FIG. B.2 – Résultats de l’algorithme de suivi SIR/modèle à vitesse constante de la figure B.1. A gauche en haut : l’erreur en position. A gauche en bas : l’erreur en vitesse. A droite l’efficacité de l’échantillonnage.

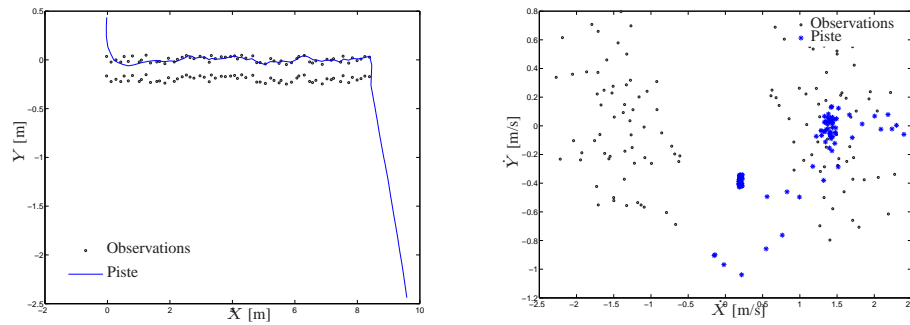


FIG. B.3 – Scénario qui propose le suivi d’un piéton sur une trajectoire complexe comportant un changement brusque de direction. A gauche, résultat du suivi en position et à droite, résultat du suivi en vitesse. Les mesures sont représentées par des cercles gris.

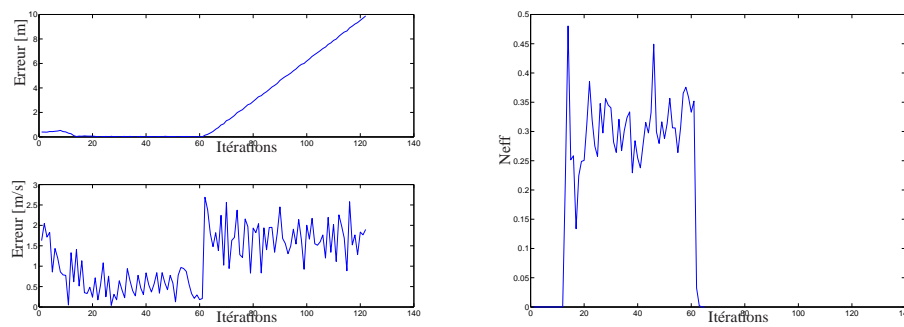


FIG. B.4 – Résultats de l’algorithme de suivi SIR/modèle à vitesse constante de la figure B.3. A gauche, l’erreur en position et en vitesse, à droite l’efficacité de l’échantillonnage.

Les figures B.1 et B.2, montre que le filtre SIR couplé à un modèle d’évolution à vitesse constante est efficace et robuste aux trajectoires linéaires.

En revanche dans le cadre d’un scénario très complexe tel qu’un demi-tour (cf. les

figures B.3 et B.4), l'algorithme est incapable de suivre cette trajectoire à cause de l'inertie de la vitesse qui empêche le modèle de prédire l'état d'un objet qui change brusquement de vitesse (passant d'une vitesse  $\dot{X} = 1.5m/s$  à une vitesse de  $\dot{X} = -1.5m/s$ ).

En ce qui concerne la reconstruction de la vitesse, ce modèle d'évolution permet d'en obtenir une bonne estimation tant que le suivi en position est correct.

Enfin, cet algorithme possède une bonne efficacité d'échantillonnage ( $N_{eff}$ ) avec une moyenne de 45 %.

### Modèle à mouvement circulaire

Pour une comparaison cohérente des deux modèles d'évolution, le filtre SIR est paramétré identiquement avec notamment  $N = 400$  particules. La dynamique est définie par le modèle à mouvement circulaire selon les équations B.4, B.5 et B.3 avec  $\sigma_x = \sigma_y = 2m.s^{-2}$ . Le bruit de mesure ( $R$ ) est toujours fixé à  $\sigma_{cx} = \sigma_{cy} = 0.1m$ . Enfin, le seuil de la fenêtre de validation reste fixé à  $\gamma = 5.99$  (voir table du  $\chi^2$ ).

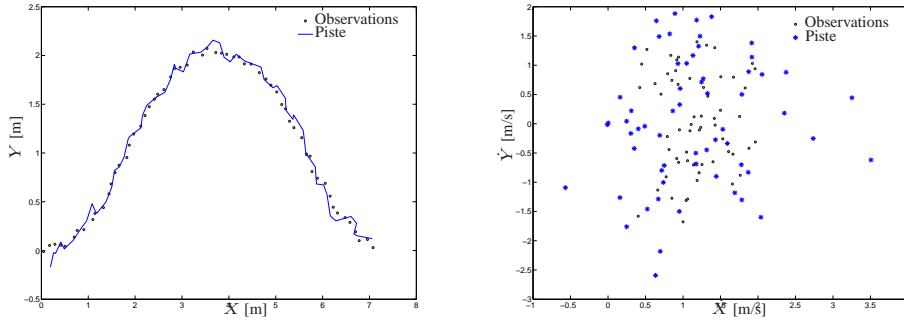


FIG. B.5 – Scénario qui propose le suivi d'un piéton sur la trajectoire en forme de courbe. A gauche, vue de dessus du résultat de suivi en position. A droite, vue de dessus du résultat de suivi en vitesse. Les mesures sont représentées par des cercles gris.

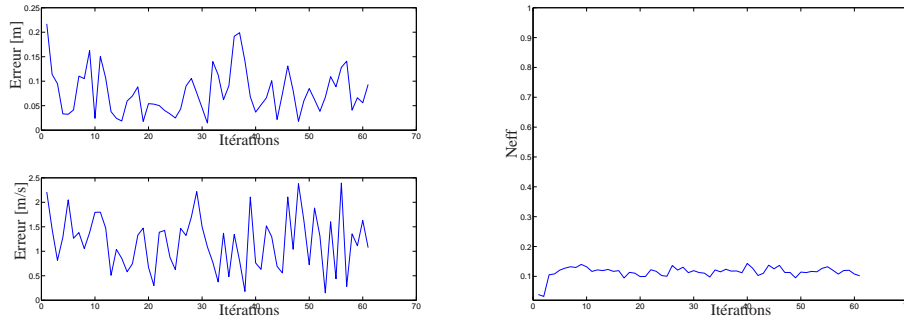


FIG. B.6 – Résultats de l'algorithme de suivi SIR/modèle à vitesse constante de la figure B.5. A gauche en haut : l'erreur en position. A gauche en bas : l'erreur en vitesse. A droite l'efficacité de l'échantillonnage.



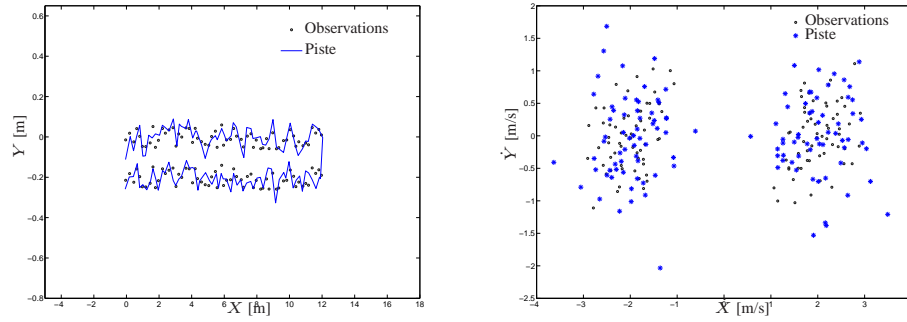


FIG. B.7 – Scénario qui propose le suivi d’un piéton sur une trajectoire complexe comportant un changement brusque de direction. A gauche, résultat du suivi en position et à droite, résultat du suivi en vitesse. Les mesures sont représentées par des cercles gris.

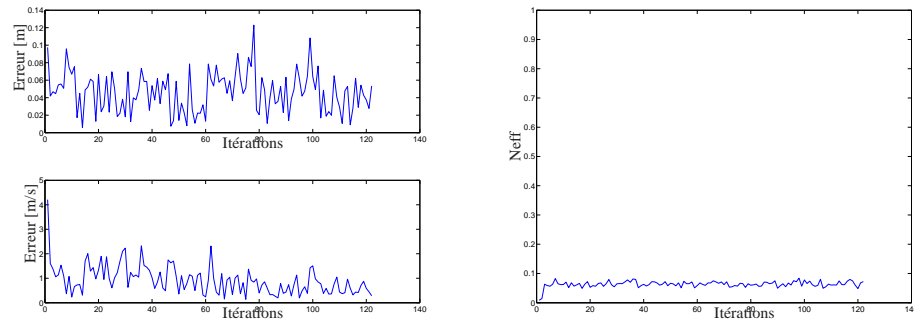


FIG. B.8 – Résultats de l’algorithme de suivi SIR/modèle à vitesse constante de la figure B.7. A gauche, l’erreur en position et en vitesse, à droite l’efficacité de l’échantillonnage.

L’algorithme SIR couplé à un modèle d’évolution à mouvement circulaire est efficace et robuste sur les deux scénarios testés. (cf. les figures B.5, B.6, B.7 et B.8).

En revanche, la reconstruction de la vitesse est moins fidèle donc moins précise. Cette estimation s’explique par un suivi en position plus oscillant.

Enfin, cet algorithme possède une mauvaise efficacité d’échantillonnage ( $N_{eff}$ ) avec une moyenne de 10 % ce qui montre une propagation plus diffuse des particules et explique en partie la robustesse du suivi en position.

## B.3 Conclusion

Le modèle d’évolution doit permettre de limiter la propagation des particules à l’étendue que peuvent physiquement atteindre les objets suivis. Si l’objet est supposé un piéton, sa dynamique peu comporter des changements brusques de direction ou de rythme. Les simulations présentées dans cette annexe ont montré l’intérêt d’utiliser un modèle à mouvement circulaire afin de pouvoir suivre plus efficacement des trajectoires complexes telles que celles des piétons.



---

## Annexe C

# Comparaison des différents algorithmes de détection laser livrés par les partenaires du projet LOVE

Cette annexe propose de comparer les performances des algorithmes de détection de piéton développés lors du projet LOVE par les différents partenaires. Ceci permet de situer les performances de l’algorithme de détection présentées dans cette thèse avec d’autres approches. Dans le contexte du projet LOVE, chaque module peut-être évalué suivant deux critères :

- ▷ détection de zones d’intérêts,
- ▷ détection et reconnaissance de piétons,

Quatre laboratoires ont proposé un algorithme fondé sur l’utilisation d’un capteur laser :

- ▷ CAO24A : module de détection de zones d’intérêts et de reconnaissance de piétons fournie par le CAOR <sup>1</sup>,
- ▷ LAS24A : module de détection de zones d’intérêts et de reconnaissance de piétons fournie par le LASMEA et présenté dans cette thèse,
- ▷ LIV24A : module de détection de zones d’intérêts fournie par le LIVIC <sup>2</sup>,
- ▷ HEU24A : module de détection de zones d’intérêts et de reconnaissance de piétons fournie par l’HeuDiaSyC <sup>3</sup>.

### C.1 Evaluation quantitative des modules définissant des détections de zones d’intérêts

Pour chaque zone d’intérêt, on mesure une distance  $d$  entre le centre du bord inférieur de la zone d’intérêt et le point au sol donné par la vérité terrain. Ces zones se divisent en « vrai positif » et « faux positif » suivant si la distance  $d$  est inférieur ou supérieur à la largeur de la détection fournie par les modules testés.

Chaque zone d’intérêt est retournée avec un indice de confiance  $I_c^Z$  entre 0 et 10000 (confiance maximale). Pour chaque indice de confiance  $I_c^Z$ , on considère toutes les zones d’intérêt d’indice égal ou supérieur (quand  $I_c^Z = 0$ , toutes les zones sont considérées ;

---

<sup>1</sup>Centre de robotique de l’école des mines de Paris

<sup>2</sup>Laboratoire sur les Interactions Véhicules-Infrastructures-Conducteurs

<sup>3</sup>Heuristique et Diagnostic des systèmes Complexes de l’université de technologie de Compiègne

quand  $I_c^Z = 10000$ , aucune zone n'est considérée).

Pour chaque indice de confiance  $I_c^Z$  on définit :

▷ le taux de bonnes détections :

$$T_{BD}^Z(I_c^Z) = \frac{N_{Z\_vrai\_positif}(I_c^Z)}{N_{pietons\_VT}} * 100 \quad (C.1)$$

où  $N_{Z\_vrai\_positif}(I_c^Z)$  correspond au nombre de zones étiquetées « vrai positif » pour l'indice de confiance ( $I_c^Z$ ) et  $N_{pietons\_VT}$  correspond au nombre de piétons de la vérité de terrain.

▷ le taux de bonne sélectivité :

$$T_S^Z(I_c^Z) = \frac{N_{Z\_vrai\_positif}(I_c^Z)}{N_{Z\_vrai\_positif}(I_c^Z) + N_{Z\_faux\_positifs}(I_c^Z)} * 100 \quad (C.2)$$

où  $N_{Z\_faux\_positifs}(I_c^Z)$  correspond au nombre de zones étiquetées « faux positif » pour l'indice de confiance ( $I_c^Z$ ).

La courbe (cf. la figure C.1 groupant les points  $[T_{BD}^Z(I_c^Z); T_S^Z(I_c^Z)]$ ) est caractéristique de l'efficacité de l'algorithme. Un module est considéré comme performant si l'indice  $I_{c\_ref}^Z$  est dans la zone délimitée par des pointillés située au plus près du point idéal  $[100\%; 100\%]$ .

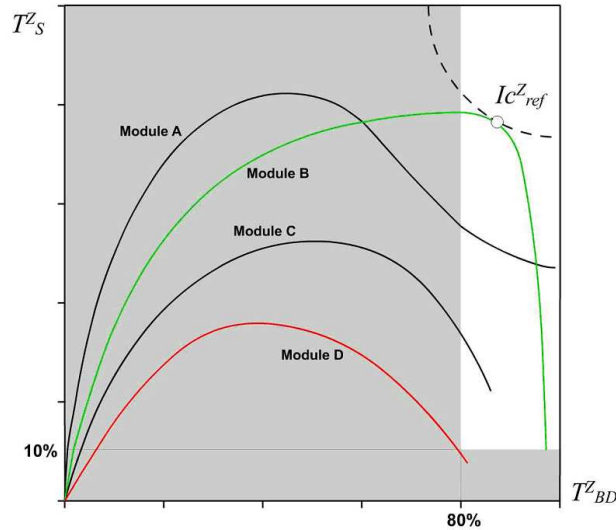


FIG. C.1 – Courbe de caractérisation des performances de détection des zones d'intérêts.

## C.2 Evaluation quantitative des modules de reconnaissance de piétons

Les modules de reconnaissance de piétons sont évalués sur la base des zones d'intérêts fournies par le module de détection de zones d'intérêts considérée(s). Les algorithmes de reconnaissance fournissent un piéton associé à un indice de confiance  $I_c^P$  compris entre 0 (objet non reconnu comme piéton) et 10000 (objet absolument reconnu comme piéton). Ces piétons se divisent en  $N_{P\_vrai\_positif}(I_c^P)$  vrais positifs et  $N_{P\_faux\_positif}(I_c^P)$  faux positifs selon qu'ils correspondent ou non à des piétons de la vérité de terrain.

Pour chaque indice de confiance  $I_c^P$  on définit :

▷ le taux de bonnes détections :

$$T_{vrai\_positif}^P(I_c^P) = \frac{N_{P\_vrai\_positif}(I_c^P)}{N_{Z\_vrai\_positif}} * 100 \quad (C.3)$$

▷ le taux de fausses détections :

$$T_{faux\_positif}^P(I_c^P) = \frac{N_{P\_faux\_positif}(I_c^P)}{N_{P\_faux\_positif}(I_c^P) + N_{Z\_faux\_positifs}(I_c^Z)} * 100 \quad (C.4)$$

La courbe (cf. la figure C.2 groupant les points  $[T_{faux\_positif}^P(I_c^P); T_{vrai\_positif}^P(I_c^P)]$  est caractéristique de l'efficacité de discrimination de l'algorithme. Un module est considéré comme performant si l'indice  $I_{c\_ref}^P$  est dans la zone délimitée par des pointillés située au plus près du point idéal  $[0\%; 100\%]$ .

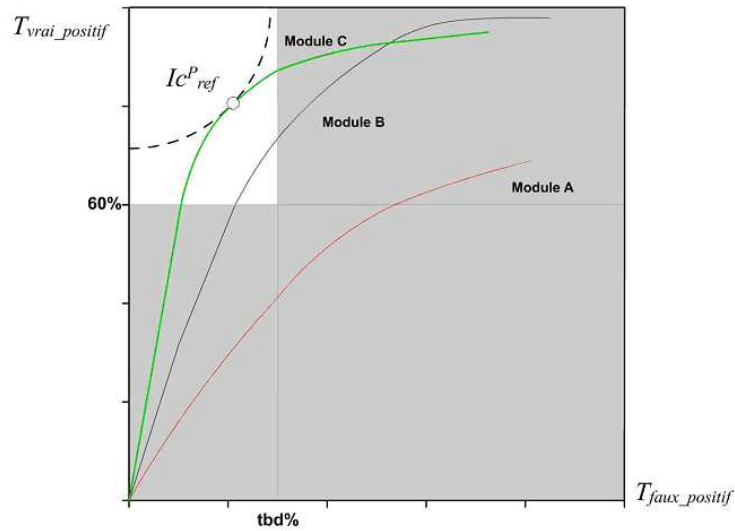


FIG. C.2 – Courbe ROC de caractérisation des performances en reconnaissance.

## C.3 Résultats

A l'heure de ces expérimentations, trois laboratoires ont livré leurs modules. Une comparaison entre ces trois modules (CAO24A, LAS24A et LIV24A) est donc proposée.

### C.3.1 Taux de détections de zones d'intérêts

La figure C.3 illustre le taux de sélectivité des trois algorithmes en fonction du nombre de zones d'intérêts détectées sur quatre scénarios choisis par l'équipementier automobile Valeo. Les fonctionnalités suivantes sont évaluées :

- ▷ taux de sélectivité,
- ▷ taux de détection des « vrais-positifs », c'est-à-dire des zones appartenant à la vérité terrain,
- ▷ nombre total de détections.

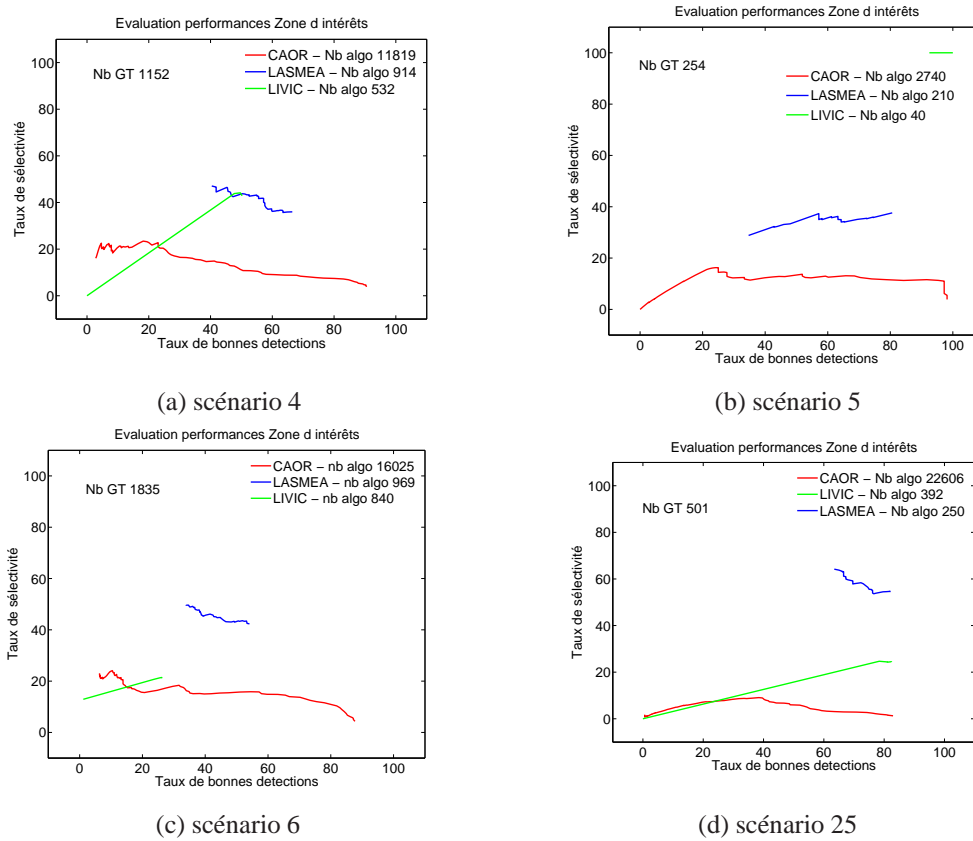


FIG. C.3 – Taux de sélectivité en fonction du nombre de zones d'intérêts détectées sur quatre scénarios Renault. Le nombre de zone d'intérêts « vérité terrain » est précisé par la variable « Nb GT ». Le nombre total de détections est donné par la variable « Nb algo ».

D'un point de vue du nombre total de détections, le module du CAOR détecte beaucoup de zones d'intérêts avec un nombre total de détections très supérieur au nombre de zones de la vérité terrain. Cette forte détection permet au CAOR de toujours dépasser les 80% de zones d'intérêts appartenant à la vérité terrain. Ce module montre cependant un taux de sélectivité relativement bas ( $< 20\%$ ).

Contrairement au CAOR, l'algorithme proposé par le LASMEA a choisi de se focaliser sur les groupes de points présentant une « signature piéton » dans l'image laser. Le nombre total de détections des zones d'intérêts est donc plus raisonnable mais diminue le taux de bonnes détections entre 60% et 80%. En revanche, cette détection plus ciblée permet d'augmenter sensiblement le taux de sélectivité qui avoisine les 50%.

Enfin, l'algorithme du LIVIC présente également un nombre total de détections raisonnable sans permettre d'obtenir une amélioration du taux de bonnes détections ou de sélectivité.

### C.3.2 Taux de reconnaissance des piétons

La figure C.4 illustre la capacité des algorithmes à discriminer les zones d'intérêts appartenant à des piétons ou non dans les quatre mêmes scénarios. Les fonctionnalités suivantes sont évaluées :

- ▷ taux de bonnes détections,

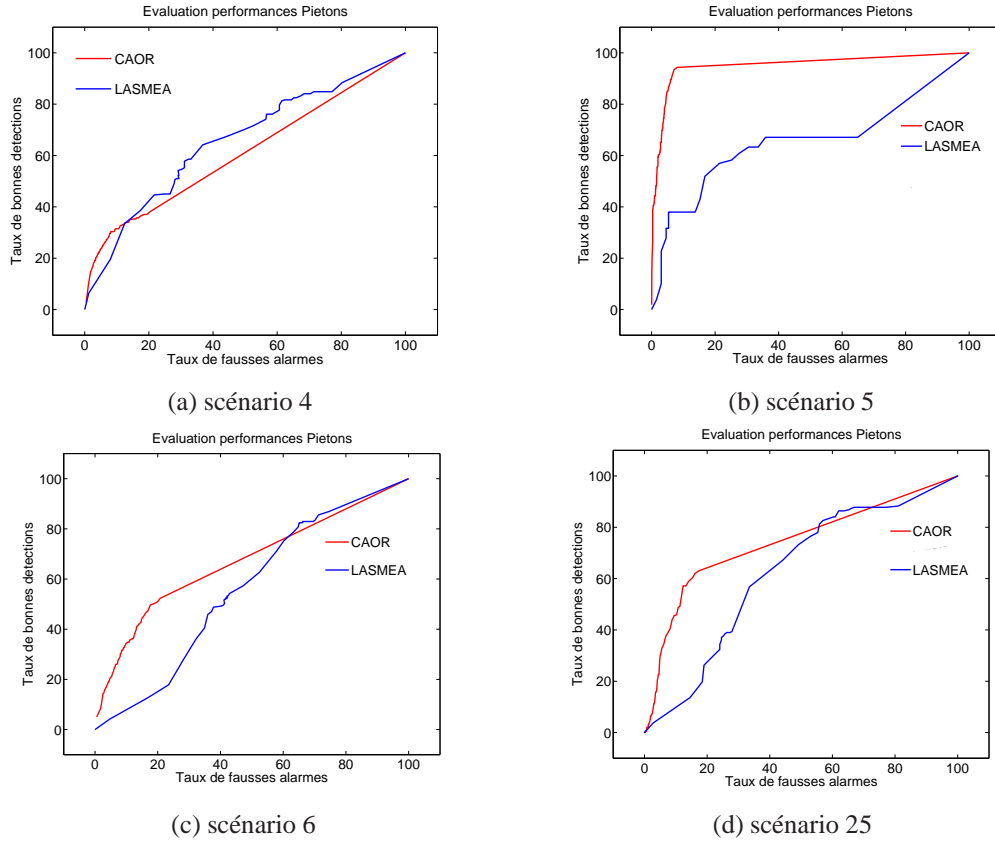


FIG. C.4 – Courbes ROC correspondant à la reconnaissance des piétons.

▷ taux de fausses alarmes.

D'un point de vue reconnaissance, seul le CAOR et le LASMEA proposent un indice de confiance ( $I_c^P$ ) permettant de construire la courbe ROC. Les résultats montrent une meilleure discrimination de l'algorithme du CAOR dans trois cas sur quatre.





---

# Bibliographie

- [obs, 1991] 1991. Observatoire national interministériel de sécurité routière. In : *Bulletin d'analyse des accidents corporels de la circulation routière*.
- [Arras, 2003] Arras, Kai Olivier. 2003. *Feature-based Robot Navigation in known and unknown Environments*. Ph.D. thesis, Ecole polytechnique fédérale de Lausanne.
- [Arras & Vestli, 1998] Arras, K.O., & Vestli, S.J. 1998. Hybrid, High-Precision Localisation for the Mail Distributing Mobile Robot System MOPS. In : *IEEE International Conference on Robotics and Automation (ICRA)*.
- [Arras et al., 2007] Arras, K.O., Mozos, O.M., & Burgard, W. 2007. Using Boosted Features for the Detection of People in 2D Range Data. In : *IEEE International Conference on Robotics and Automation (ICRA)*.
- [Athans et al., 1968] Athans, M., Wishner, R.P., & Bertolini, A. 1968. Suboptimal state estimate for continuous-time nonlinear systems from discrete noisy measurements. *IEEE Trans Automatica control*, **13**(5), 504–514.
- [Baker & Matthews, 2004] Baker, S., & Matthews, I. 2004. Lucas-Kanade 20 Years On : A Unifying Framework. *International Journal of Computer Vision*, **56**(3), 221–255.
- [Bar-Shalom & Fortmann, 1988] Bar-Shalom, Y., & Fortmann, T.E. 1988. *Alignment and Data Association*. New-York : Academic.
- [Bar-Shalom & Li, 1995] Bar-Shalom, Y., & Li, X.R. 1995. *Multitarget-Multisensor Tracking : Principles and Techniques*. Storrs, CT : YBS Publishing.
- [Bardet, 2009] Bardet, François. 2009. *Suivi et Catégorisation Multi-Objets par Vision Artificielle : Applications au suivi de Personnes et de Véhicules*. Ph.D. thesis, Université Blaise Pascal.
- [Bertozzi et al., 2007] Bertozzi, M., Broggi, A., Rose, M. Del, & Felisa, M. 2007. A Symmetry-based Validator and Refinement System for Pedestrian Detection in Far Infrared Images. In : *International Transport System Conference (ITSC)*.
- [Besl & McKay, 1992] Besl, P.J., & McKay, N.D. 1992. A method for registration of 3-D shapes. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **14**(2), 239–256.
- [Blanc, 2005] Blanc, Christophe. 2005. *Combinaison d'estimations : Application à la détection d'obstacles à bord des véhicules routiers intelligents*. Ph.D. thesis, Université Blaise Pascal.
- [Borges & Aldon, 2004] Borges, G.A., & Aldon, M.J. 2004. Line Extraction in 2D Range Images for Mobile Robotics. *Journal of intelligent and Robotic Systems*, **40**, 267–297.
- [Bréhard & Cadre, 2004] Bréhard, T., & Cadre, J-P. Le. 2004. A new approach for the bearings-only problem : estimation of the variance-to-range ratio. In : *International Conference on Information Fusion*.

- [Broggi *et al.*, 2008] Broggi, Alberto, Cerri, Pietro, Ghidoni, Stefano, Grisleri, Paolo, & Jung, Ho Gi. 2008. Localization and Analysis of Critical Areas in Urban Scenarios. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Chanier *et al.*, 2008] Chanier, F., Checchin, P., Blanc, C., & Trassoudaine, L. 2008. Map fusion based on a multi-map SLAM framework. *In : IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*.
- [Chateau, 1999] Chateau, Thierry. 1999. *Détection robuste d'interface par fusion d'informations incertaines : application à l'aide au guidage d'engin agricoles*. Ph.D. thesis, Université Blaise Pascal.
- [Checchin, 1996] Checchin, Paul. 1996. *Segmentation d'images de profondeur*. Ph.D. thesis, Université Blaise Pascal.
- [Cornou *et al.*, 2008] Cornou, S., Fleury, B., & Trassoudaine, L. 2008. Pedestrian accident context and technologic development in LOVe project. *In : V.I.S.I.O.N, Vehicle and Infrastructure Safety Improvement in Adverse Conditions and Night Driving*.
- [Cox, 1991] Cox, I.J. 1991. Blanche : An Experiment in Guidance and Navigation of an Autonomous Robot Vehicle. *IEEE Transactions on Robotics and Automation*, **7**, 193–204.
- [Cui *et al.*, 2005] Cui, J., Zha, H., Zhao, H., & Shibasaki, R. 2005. Tracking Multiple People using Laser and Vision. *In : IEEE Intelligent Robots and Systems (IROS)*.
- [Cui *et al.*, 2007] Cui, J., Zha, H., Zhao, H., & Shibasaki, R. 2007. Laser-based detection and tracking of multiple people in crowds. *Comput. Vis. Image Underst.*, **106**, 300–312.
- [Denoeux & Govaert, 1996] Denoeux, T., & Govaert, G. 1996. Combined supervised and unsupervised learning for system diagnosis using Dempster-Shafer theory. *In : Conference Computational Engineering Applications (CESA)*.
- [Diosi & Kleeman, 2003] Diosi, A., & Kleeman, L. 2003. Uncertainty of Line Segments Extracted from Statistic SICK PLS Laser Scans. *In : Australian Conference on Robotics and Automation (ACRA)*.
- [Douillard *et al.*, 2007] Douillard, B., Fox, D., & Ramos, F. 2007. A spatio-temporal probabilistic model for multi-sensor object recognition. *In : IEEE Intelligent Robots and Systems (IROS)*.
- [Egammal *et al.*, 2000] Egammal, A., Harwood, D., & Davis, L.S. 2000. Non-parametric Model for Background Subtraction. *In : European Conference on Computer Vision (ECCV)*.
- [Elfes, 1989] Elfes, Alberto. 1989. *Occupancy grids : A probabilistic framework for robot perception and navigation*. Ph.D. thesis, Electrical and Computer Engineering Departement/Robotics Institute, Carnegie Mellon University.
- [Fayad *et al.*, 2008] Fayad, F., Cherfaoui, V., & Dherbomez, G. 2008. Updating confidence indicators in a multi-sensor pedestrian tracking system. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Fiscler & Bolles, 1981] Fiscler, M., & Bolles, R. 1981. Random sample consensus : A paradigm for model fitting with application to image analysis and automated cartography. *Communication of the ACM*, **24**, 381–395.
- [Fod *et al.*, 2002] Fod, A., Howard, A., & Matarié, M.J. 2002. Laser-Based People Tracking. *In : IEEE International Conference on Robotics and Automation (ICRA)*.

- [Fontaine, 1995] Fontaine, H. 1995. A Typological Analysis of Pedestrian Accidents. *In : 7th workshop of ICTCT*.
- [Fournier *et al.*, 2006] Fournier, R.A., Leblanc, S.G., Langelier, F., & Côté, J.F. 2006. Utilisation du laser imageur 3D pour mesures en forêt. *In : Atelier REGLIS*.
- [Franck *et al.*, 2003] Franck, O., Nieto, J., Guivant, J., & Scheduling, S. 2003. Multiple target tracking using sequential Monte Carlo methods and statistical data association. *In : IEEE Intelligent Robots and Systems (IROS)*.
- [Freitas *et al.*, 2001] Freitas, N., Andrieu, C., Sorensen, P.H., Niranjana, M., & Gee, A. 2001. *Sequential Monte Carlo methods for neural networks*. Springer-Verlag.
- [Fuerstenberg & Dietmayer, 2004] Fuerstenberg, K.C., & Dietmayer, K. 2004. Object Tracking and Classification for Multiple Active Safety and Comfort Applications using Multilayer Laserscanner. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Fuerstenberg & Lages, 2005] Fuerstenberg, K.C., & Lages, U. 2005. New European Approach for Intersection Safety - The EC-Project INTERSAFE. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Fuerstenberg *et al.*, 2002] Fuerstenberg, K.C., Dietmayer, K.C.J., & Willhoeft, V. 2002. Pedestrian Recognition in Urban Traffic using a vehicle based Multilayer Laserscanner. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Fuerstenberg *et al.*, 2003] Fuerstenberg, K.C., Linzmeier, D.T., & Dietmayer, K.C.J. 2003. Pedestrian Recognition and Tracking of Vehicles using a vehicle based Multilayer Laserscanner. *In : 10th World Congress on Intelligent Transport Systems (ITS)*.
- [Gallice, 2002] Gallice, J. 2002. *Images de profondeur, Traité IC2 Information Commande Communication, Traitement du Signal et de l'Image*. Hermès Science Publications.
- [Gate & Nashashibi, 2008] Gate, Gwennael, & Nashashibi, Fawzi. 2008. Using Targets Appearance to Improve Pedestrian Classification with a Laser Scanner. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Gavrila *et al.*, 2001] Gavrila, D., Kunert, M., & Lages, U. 2001. A multi-sensor approach for the protection of vulnerable traffic participants-the PROTECTOR project. *IEEE Instrumentation and Measurement Technology Conference*, 2044–2048.
- [Gavrila, 1999] Gavrila, D.M. 1999. The visual analysis of human movement : A survey. *Computer Vision and Image Understanding*, **72(1)**, 82–98.
- [Gavrila, 2001] Gavrila, D.M. 2001. Sensor-based pedestrian protection. *IEEE Intelligent Systems*, **16**, 77–81.
- [Gidel *et al.*, 2008a] Gidel, S., Checchin, P., Blanc, C., Chateau, T., & Trassoudaine, L. 2008a. Détection de piéton à l'aide d'un capteur laser quads embarqué. *In : 16ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, RFIA*.
- [Gidel *et al.*, 2008b] Gidel, S., Blanc, C., Chateau, T., Checchin, P., & Trassoudaine, L. 2008b. Non-Parametric Data Association for Particle Filter based Multi-Object Tracking. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Gidel *et al.*, 2009a] Gidel, S., Blanc, C., Chateau, T., Checchin, P., & Trassoudaine, L. 2009a. A Method based Multilayer Laserscanner to Detect and Track Pedestrian in Urban Environment. *In : IEEE Intelligent Vehicles Symposium (IV)*.

- [Gidel *et al.*, 2009b] Gidel, S., Blanc, C., Chateau, T., Checchin, P., & Trassoudaine, L. 2009b. Non-parametric Laser and Video Data Fusion : Application to Pedestrian Detection in Urban Environment. *In : 12th International Conference on Information Fusion*.
- [Gordon *et al.*, 1993a] Gordon, N., Salmond, D., & Smith, A. 1993a. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE Proc. F, Radar and Signal Processing*, **140(2)**, 107–113.
- [Gordon *et al.*, 1993b] Gordon, N.J., Salmond, D.J., & Smith, A.F.M. 1993b. Novel approach to non-linear/non-Gaussian Bayesian state estimation. *Radar and Signal Processing*, **140**, 107–113.
- [Gorji *et al.*, 2007] Gorji, A., Shiry, S., & Menhaj, B. 2007. Multiple Target Tracking For Mobile Robots using the JPDAF Algorithm. *In : IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*.
- [Goya *et al.*, 2008] Goya, Y., Chateau, T., Malaterre, L., & Trassoudaine, L. 2008. Estimation Précise de la Trajectoire d'un Véhicule par Fusion de Capteurs Vision et télémètre Laser. *In : 16ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle (RFIA)*.
- [Grenier, 2005] Grenier, Thomas. 2005. *Apport de l'espace des caractéristiques et des paramètres d'échelle adaptatifs pour le filtrage et la segmentation d'image*. Ph.D. thesis, Institut National des Sciences Appliquées de Lyon.
- [Gustafsson *et al.*, 2002] Gustafsson, F., Gunnarson, F., Bergman, N., U.Forssell, Jansson, J., Karlsson, R., & Nordlund, P-J. 2002. Particle filters for positioning, navigation and tracking. *IEEE Transactions on Signal Processing*, **50(2)**, 425–437.
- [Gutmann & Konolige, 1994] Gutmann, J.S., & Konolige, K. 1994. Robot pose estimation in unknown environments by matching 2d range scans. *In : IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [Gutmann & Konolige, 1999] Gutmann, J.S., & Konolige, K. 1999. Incremental mapping of large cyclic environments. *In : IEEE International Symposium on Computational Intelligence in Robotics and Automation*.
- [Gutmann & Schlegel, 1996] Gutmann, J.S., & Schlegel, C. 1996. AMOS : Comparaison of Scan Matching Approaches for Self-Localization in Indoor Environment. *In : 1st Euromicro Workshop on Advanced Mobile Robots*.
- [Hastings, 1970] Hastings, W. 1970. Monte Carlo Sampling Methods Using Markov Chains and their Applications. *Biometrika*, **57(1)**, 97–109.
- [Hue *et al.*, 2002] Hue, C., Cadre, J-P. Le, & Pérez, P. 2002. Tracking multiple objects with particle filtering. *IEEE Transactions on Aerospace and Electronic Systems*, **38(3)**, 791–812.
- [Isard & Blake, 1996] Isard, M., & Blake, A. 1996. Contour tracking by stochastic propagation of conditional density. *Lecture Notes in Computer Science*, **1064**, 343–356.
- [Isard & Blake, 1998] Isard, M., & Blake, A. 1998. CONDENSATION - Conditional density propagation for visual tracking. *Int. J. Comput*, **29**, 5–28.
- [Jouannin, 1999] Jouannin, Stéphane. 1999. *Association et fusion de données : application au suivi et à la localisation d'obstacles par radar à bord d'un véhicule routier intelligent*. Ph.D. thesis, Université Blaise Pascal, France.



- [Julier & Uhlmann, 1997] Julier, S., & Uhlmann, J. 1997. A new extension of the Kalman filter to nonlinear systems. *In : International Symposium Aerospace/Defense Sensing, Simulations and Controls*.
- [Kalman, 1960] Kalman, R. 1960. A new approach to linear filtering and prediction problems. *Transactions of the ASME Journal of Basic Engineering*, **82**, 35–45.
- [Kalman & Bucy, 1961] Kalman, R., & Bucy, R. 1961. New results in linear filtering and prediction theory. *Transactions of the ASME Journal of Basic Engineering*, **83**, 93–108.
- [Karlsson & Gustafsson, 2001] Karlsson, R., & Gustafsson, F. 2001. Monte Carlo data association for multiple target tracking. *In : IEEE Target tracking : Algorithms and applications*.
- [Khan *et al.*, 2004] Khan, Z., Balch, T., & Dellaert, F. 2004. An MCMC-based particle filter for tracking multiple interacting targets. *In : Proc. European Conf. on Computer Vision (ECCV)*.
- [Kitagawa, 1996] Kitagawa, G. 1996. Monte Carlo filter and smoother form non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, **5**(1), 1–25.
- [Kluge *et al.*, 2001] Kluge, B., Koehler, C., & Prassler, E. 2001. Fast and Robust Tracking of Multiple Moving Objects with a Laser Range Finder. *In : IEEE International Conference on Robotics and Automation (ICRA)*.
- [Kurazume *et al.*, 2008] Kurazume, R., Yamada, H., Murakami, K., & ans T. Hasegawa, Y. Iwashita. 2008. Target Tracking Using SIR and MCMC Particle Filters by Multiple Cameras and Laser Range Finders. *In : IEEE Intelligent Robots and Systems (IROS)*.
- [Lefevre *et al.*, 2004] Lefevre, E., Manata, J.P., & Jolly, D. 2004. Classification par la théorie de l'évidence pour la gestion de tournée de véhicules. *In : 14ème Congrès Francophone Reconnaissance des Formes et Intelligence Artificielle (RFIA)*.
- [Li & Jilkov, 2000] Li, X.R., & Jilkov, V.P. 2000. A Survey of Maneuvering Target Tracking : Dynamic Models. *In : Proc. of SPIE Conference on Signal and Data Processing of Small Target*.
- [Lindstrom & Eklundh, 2001] Lindstrom, M., & Eklundh, J.O. 2001. Detecting and tracking moving objects from a mobile platform using a laser range scanner. *In : IEEE Intelligent Robots and Systems (IROS)*.
- [Lu & Milios, 1997] Lu, F., & Milios, E. 1997. Robot pose estimation in unknown environment mapping. *Autonomous Robots*, **18**, 249–275.
- [Meinecke *et al.*, 2005] Meinecke, M.M., Obojski, M.A., Tons, M., & Dehesa, M. 2005. SAVE-U : First experiences with a pre-crash system for enhancing pedestrian safety. *In : Proc. 5th Eur. Congr. Intell. Transp.*
- [Mendes *et al.*, 2004] Mendes, A., Bento, L.C., & Nunes, U. 2004. Multi-target Detection and Tracking with a Laserscanner. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Mochon & McMahon, 1980] Mochon, S., & McMahon, T.A. 1980. Ballistic Walking. *Journal of Biomechanics*, **13**, 49–57.
- [Monteiro *et al.*, 2006] Monteiro, G., Premevida, C., Peixoto, P., & Nunes, U. 2006. Tracking and Classification of Dynamic Obstacles Using Laser Range Finder and Vision. *In : IEEE Intelligent Robots and Systems (IROS) Workshop : Safe Navigation in Open and Dynamic Environments - Autonomous Systems versus Driving Assistance Systems*.

- [Montemerlo *et al.*, 2002] Montemerlo, M., Thrun, S., & Whittaker, W. 2002. Conditional particle filters for simultaneous mobile robot localization and people tracking. In : *IEEE International Conference on Robotics and Automation (ICRA)*.
- [Moral *et al.*, 1992] Moral, P. Del, Rigal, G., & Salut, G. 1992. *Estimation et commande optimale non linéaire : un cadre unifié pour la résolution particulière*. Rapport de Recherche n°2, Convention DRET-DIGILOG-LAAS/CNRS, LAAS, 1992.
- [Nabaa & Bishop, 1999] Nabaa, N., & Bishop, R. 1999. Solution to a multisensor tracking problem with sensor registration errors. *IEEE Trans. Aerospace and Electronic Systems*, **35**, 354–362.
- [Navarro-Serment *et al.*, 2008] Navarro-Serment, L.E., Mertz, C., Vandapel, N., & Herbert, M. 2008. LADAR-based Pedestrian Detection and Tracking. In : *IEEE International Conference on Robotics and Automation (ICRA) Workshop : Human Detection from Mobile Robot Platforms : Different Perspective, Different Modalities*.
- [Ngako-Pangop *et al.*, 2008] Ngako-Pangop, L., Chausse, F., Cornou, S., & Chapuis, R. 2008. Asynchronous Bayesian Algorithm for Object Classification : Application to Pedestrian Detection in Urban Areas. In : *11th International Conference on Information Fusion*.
- [Nguyen *et al.*, 2005] Nguyen, Viet, Martinelli, Agostino, Tomasis, Nicolas, & Siegwart, Roland. 2005. A Comparison of Line Extraction Algorithms using 2D Laser Range-finder for Indoor Mobile Robotics. In : *IEEE Intelligent Robots and Systems (IROS)*.
- [Oh *et al.*, 2004] Oh, S., Russell, S., & Sastry, S. 2004. Markov chain Monte Carlo Data Association for Multiple-Target Tracking. In : *IEEE Conference on Decision and Control*.
- [Okuma *et al.*, 2004] Okuma, K., Taleghani, A., Freitas, O., Little, J.J., & Lowe, D.G. 2004. A boosted particle filter : Multitarget detection and tracking. In : *Proc. European Conf. on Computer Vision (ECCV)*.
- [Oliveira, 1994] Oliveira, P. Milheiro. 1994. Approximate filters for a nonlinear discrete time filtering problem with small observation noise. *Stochastics and Stochastics Reports*, **46(24)**, 1–24.
- [Paden *et al.*, 2004] Paden, M., Scurfiel, R., Sleet, D., D. Mohan, A.A. Hyder, Jarawan, E., & Mathers, C. 2004. World report on road traffic injury prevention. In : *World Health Organization (WHO)*.
- [Papageorgiou & Poggio, 2000] Papageorgiou, C., & Poggio, T. 2000. A trainable system for object detection. *Int. Journal of Computer Vision*, **38**, 15–33.
- [Picard, 1991] Picard, J. 1991. Efficiency of the extended kalman filter for nonlinear systems with small noise. *Journal on Applied Mathematics*, **51(3)**, 843–885.
- [Prassler *et al.*, 1999] Prassler, E., Scholz, J., & Fiorini, P. 1999. Navigating a Robotic Wheelchair in a Railway Station during Rush Hour. *Int. Journal on Robotics Research (IJRR)*, **18**, 760–772.
- [Premebida & Nunes, 2005] Premebida, C., & Nunes, U. 2005. Segmentation and Geometric Primitives Extraction from 2D Laser Range Data for Mobile Robot Application. In : *Proc. 5th National Festival of Robotics, Scientific Meeting (ROBOTICA)*.
- [Premibida *et al.*, 2007] Premibida, C., Monteiro, G., Nunes, U., & Peixoto, P. 2007. A LIDAR and vision-based approach for pedestrian and vehicle detection and tracking. In : *International Transport System Conference (ITSC)*.



- [Ramer, 1972] Ramer, U. 1972. An Iterative Procedure for the Polygonal Approximation of Plane Curves. *Computer Graphics and Image Processing*, **1**, 244–256.
- [Read, 1979] Read, D. 1979. An algorithm for tracking multiple targets. *IEEE Transactions on Automation and Control*, **24**, 84–90.
- [Sarkka *et al.*, 2004] Sarkka, S., Vehtari, A., & Lampinen, J. 2004. Rao-Blackwellized Particle Filter for Multiple Target Tracking. *In : 7th International Conference on Information Fusion*.
- [Schulz *et al.*, 2001] Schulz, D., Burgard, W., Fox, D., & Cremers, A.B. 2001. Tracking Multiple Moving Objects with a Mobile Robot. *In : IEEE Computer Vision and Pattern Recognition (CVPR)*.
- [Shao *et al.*, 2007] Shao, X., Zhao, H., Nakamura, K., Katabira, K., Shibasaki, R., & Nakagawa, Y. 2007. Detection and Tracking of Multiple Pedestrians by Using Laser Range Scanner. *In : IEEE Intelligent Robots and Systems (IROS)*.
- [Siadat *et al.*, 1997] Siadat, A., Kaske, A., Klausmann, S., Dufaut, M., & Husson, R. 1997. An Optimized Segmentation Method for a 2D Laser-Scanner Applied to Mobile Robot Navigation. *In : 3rd IFAC Symposium on Intelligent Components and Instruments for Control Applications*.
- [Smith, 2007] Smith, Kevin. 2007. *Bayesian Methods for Visual Multi-Object Tracking with Applications to Human Activity Recognition*. Ph.D. thesis, Ecole Polytechnique Federale de Lausanne.
- [Spinello *et al.*, 2008] Spinello, L., Triebel, R., & Siegwart, R. 2008. Multimodal Detection and Tracking of Pedestrians in Urban Environments with Explicit Ground Plane Extraction. *In : IEEE Intelligent Robots and Systems (IROS)*.
- [Spinello *et al.*, 2009] Spinello, L., Triebel, R., & Siegwart, R. 2009. Multiclass Multimodal Detection and Tracking in Urban Environments. *In : 7th Field and Service Robotics (FSR)*.
- [Szarvas *et al.*, 2006] Szarvas, M., Sakai, U., & Ogata, J. 2006. Real-time Pedestrian Detection using LIDAR and Convolutional Neural Network. *In : IEEE Intelligent Vehicles Symposium (IV)*.
- [Taylor & Probert, 1996] Taylor, R., & Probert, P. 1996. Range Finding and Feature Extraction by Segmentation of Images for Mobile Robot Navigation. *In : IEEE International Conference on Robotics and Automation (ICRA)*.
- [Teknomo, 2002] Teknomo, Kardi. 2002. *Microscopic pedestrian flow characteristics : Development of an image processing data collection and simulation model*. Ph.D. thesis, Tokohu University, Japan.
- [Thornton *et al.*, 2008] Thornton, S.M., Hoffelder, M., & Morris, D.D. 2008. Multi-sensor Detection and Tracking of Humans for Safe Operation with Unmanned Ground Vehicles. *In : IEEE International Conference on Robotics and Automation (ICRA) Workshop : Human Detection from Mobile Robot Platforms : Different Perspective, Different Modalities*.
- [Vandorpe *et al.*, 1996] Vandorpe, J., Brussel, H.V., & Xu, H. 1996. Exact Dynamic Map Building for a Mobile Robot using Geometrical Primitives Produced by a 2D Range Finder. *In : IEEE International Conference on Robotics and Automation (ICRA)*.

- [Vermaak *et al.*, 2002] Vermaak, J., Andrieu, C., Doucet, A., & Godsill, S.J. 2002. Particle methods for Bayesian modeling and enhancement of speech signals. *IEEE Transactions on Speech and Audio Processing*, **10**(3), 173–185.
- [Vermaak *et al.*, 2005] Vermaak, J., Godsill, S.J., & Pérez, P. 2005. Monte Carlo filtering for multi-target tracking and data association. *IEEE Transactions on Aerospace and Electronic Systems*, **41**, 309–332.
- [Vestli, 1995] Vestli, S.J. 1995. *Fast, accurate and robust estimation of mobile robot position and orientation*. Ph.D. thesis, ETH Zürich Switzerland.
- [Wang, 2004] Wang, Chieh-Chih. 2004. *Simultaneous Localization, Mapping and Moving Object Tracking*. Ph.D. thesis, Robotics Institute Carnegie Mellon University.
- [Wulf & Wagner, 2003] Wulf, O., & Wagner, B. 2003. Fast 3D Scanning Methods For laser Measurement Systems. In : *14th International Conference on Control Systems and Computer Science*.
- [Xavier *et al.*, 2004] Xavier, J., Pacheco, M., Castro, D., Ruano, A., & Nunes, U. 2004. Fast Line, Arc/Circle and Leg Detection from Laser Scan Data in a Player Driver. In : *IEEE International Conference on Robotics and Automation (ICRA)*.
- [Zhang & Ghosh, 2000] Zhang, L., & Ghosh, B.K. 2000. Line segment Based Map Building and Localization Using 2D laser Rangefinder. In : *IEEE International Conference on Robotics and Automation (ICRA)*.
- [Zhao & Shibasaki, 2005] Zhao, H., & Shibasaki, R. 2005. A Novel System for Tracking Pedestrians Using Multiple Single-Row Laser-Range Scanners. *Systems, Man and Cybernetics, Part A, IEEE Transactions*, **35**, 283–291.
- [Zhao *et al.*, 2008] Zhao, H., Chiba, M., R. Shibasaki, K. Katabira, Cui, J., & Zha, H. 2008. Driving Safety and Traffic Data Collection - A Laser Scanner Based Approach. In : *IEEE Intelligent Vehicles Symposium (IV)*.

---

# Abstract

The research work presented for this thesis is based on computer vision and deals with the detection and tracking of pedestrians located on a vehicle's trajectory in an urban environment. In such a complex environment, one of the major difficulties is the ability to detect pedestrians among numerous other obstacles on the road ahead of the vehicle. Another key point is the ability to follow pedestrians so as to predict their trajectory and thus, if necessary, to avoid any collision with the vehicle. Several other constraints have to be faced in the industrial context of smart vehicles. Therefore some robust, real-time algorithms are required along with the cheapest possible sensors.

At first my work consisted in developing an appropriate method to detect, then identify and track pedestrians in an outdoor environment from a single four-plane laser sensor. Pedestrian extraction and the merging of the four laser planes are both based on a non-parametric kernel method, also called "Parzen Windows". Initially, this estimator is used to approximate the likelihood function of the impact record in the laser image according to a pedestrian's geometrical characteristics. Secondly this estimator is used to calculate the likelihood that a pedestrian should be located within the four laser planes. Finally, to best characterize the complex trajectory of a pedestrian, the tracking process is based on the traditional particle filter.

Unfortunately a pedestrian detection system which relies only on a laser sensor remains unsatisfactory as far as performance is concerned. Indeed, the inherent limitations of this sensor (no information about height, the outline or the color of the objects), as well as its sensitivity to such atmospheric conditions as rain or fog, make it necessary to resort to a multisensorial solution which allows to effectively combine the information provided by the laser and video sensors. This fusion methods is based on the development of a non-parametric method for data association, which allows to keep all the information contained in the measurements sent by the laser and video sensors.

The performance of each proposed algorithm was characterized and reviewed, using real data obtained from numerous recording on board the LASMEA and Renault test vehicle ; Renault being the French vehicle manufacturer with whom we collaborate on our ANR LOVE project.

## Keys words

Pedestrian detection, Pedestrian tracking, LIDAR, intelligent vehicle, particle filter, non-parametric methods, laser and video data fusion.

---

# Résumé

Les travaux présentés dans cette thèse ont pour cadre la vision par ordinateur et concernent la détection et le suivi de piéton se trouvant sur la trajectoire d'un véhicule routier circulant en milieu urbain. Dans ce type d'environnement complexe, une des difficultés majeures est la capacité à discerner les piétons des nombreux autres obstacles situés sur la chaussée. Un autre point essentiel est de pouvoir les suivre afin de prédire leur déplacement et ainsi le cas échéant éviter le contact avec le véhicule. D'autres contraintes s'ajoutent dans le contexte industriel des véhicules routiers intelligents. Il est nécessaire de proposer des algorithmes robustes temps réel avec des capteurs les moins chers possible.

Dans une première partie, mon travail a consisté à mettre au point une méthode de détection, d'identification puis de suivi de piéton en environnement extérieur à partir d'un unique capteur laser quatre plans. L'extraction d'« objets piétons » et la fusion des quatre plans laser reposent tous deux, sur une méthode non paramétrique à noyau aussi appelée « fenêtres de parzen ». Cet estimateur permet dans un premier temps, d'approximer la fonction de vraisemblance d'un relevé d'impacts laser en fonction des caractéristiques géométriques d'un piéton. Dans un second temps, cet estimateur est utilisé pour calculer la probabilité qu'un piéton soit situé dans les quatre plans laser. Afin de caractériser au mieux la trajectoire complexe d'un piéton, le pistage repose sur l'utilisation classique d'un filtre à particules.

Malheureusement un système de détection de piéton fondé uniquement sur un capteur laser reste insuffisant du point de vue de la performance. En effet, les limitations intrinsèques de ce capteur (absence d'information sur la hauteur, le contour, ou la couleur des objets), sa sensibilité aux conditions atmosphériques comme la pluie ou le brouillard, a conduit à réaliser une solution multisensorielle permettant de combiner judicieusement les informations fournies par un capteur laser et vidéo. Cette fusion repose sur le développement d'une méthode non paramétrique d'association de données permettant de conserver toute l'information contenue dans les mesures envoyées par les capteurs laser et vidéo.

Les performances des différents algorithmes proposés ont été caractérisées et commentées à partir de données réelles obtenues sur de multiples scénarios enregistrés à bord du véhicule expérimental du LASMEA et de Renault, avec qui nous avons collaboré dans le cadre du projet ANR LOVE.

**Mots clés :** Détection et suivi de piéton, LIDAR, véhicule intelligent, filtre à particules, méthodes non paramétriques, fusion de données laser et vidéo.